ROLE OF ORCA AND ORC IN CHROMATIN ORGANIZATION AND DNA REPLICATION

BY

SUMANPRAVA GIRI

DISSertation

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Cell and Developmental Biology in the Graduate College of the University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Supriya G. Prasanth, Chair and Director
Professor Lisa J. Stubbs
Professor Andrew S. Belmont
Professor Brian C. Freeman
Professor Taekjip Ha
**ABSTRACT**

In eukaryotes, Origin recognition complex (ORC) proteins establish the pre-replicative complex (pre-RC) at the origins and this is essential for the initiation of DNA replication. In human cells, ORC is a highly dynamic complex with many separate functions attributed to sub-complexes or individual subunits of ORC including heterochromatin organization, telomere and centromere function, centrosome duplication and cytokinesis.

Heterochromatic domains are enriched with repressive histone marks, including histone H3 lysine 9 methylation, written by lysine methyltransferases (KMTs). ORC along with the pre-RC protein Origin Recognition Complex-Associated (ORCA/LRWD1), preferentially localizes to heterochromatic regions in post-replicated cells. The role of ORCA and ORC in heterochromatin organization remained elusive. In Chapter II, I describe my efforts to understand the significance of ORCA-ORC’s association with heterochromatin. ORCA recognizes methylated H3K9 marks and interacts with repressive KMTs, including G9a/GLP and Suv39H1 in a chromatin context-dependent manner.

Single-molecule pull-down assays demonstrate that ORCA-ORC and multiple H3K9 KMTs exist in a single complex and that ORCA stabilizes H3K9 KMT complex. Cells lacking ORCA show alterations in chromatin architecture, with significantly reduced H3K9 di- and tri-methylation at specific chromatin sites. Changes in heterochromatin structure due to loss of ORCA affects replication timing, preferentially at the late-replicating regions. I demonstrate that ORCA acts as a scaffold for the establishment of H3K9 KMT complex and its association and activity at specific chromatin sites is crucial for the organization of heterochromatin structure.

Heterochromatin mostly constitutes tightly packaged DNA, decorated with repressive histone marks, including histone H3 methylated at lysine 9, histone H4 methylated at lysine 20 and histone H3 methylated at lysine 27. Each of these marks is incorporated by specific histone lysine methyl transferases. While constitutive heterochromatin enriched with H3K9me3 and H4K20me3 occur within repetitive elements, including centromeres and telomeres, the facultative heterochromatin resides on the inactive X-chromosome and contains H3K27me3 mark. ORCA associates with constitutive and facultative heterochromatin in human cells and binds to repressive histone marks. In Chapter III, I show that ORCA binds to multiple repressive histone methyl transferases including G9a, GLP, Suv39h1 (H3K9me2/3), Suv420h1/h2 (H4K20me2/3) and EZH2 (H3K27me3). Removal of ORCA from human cells causes aberrations in the chromatin architecture. I therefore propose that
ORCA acts as a scaffold protein that enables the formation of multiple histone lysine methyltransferase complexes at heterochromatic sites thereby facilitating chromatin organization. Open chromatin structures regulate the efficiency of preRC formation and replication initiation. However, the molecular mechanisms that affect chromatin structure and how the preRC components establish themselves on the chromatin remain to be understood. In Chapter IV, I show that human Orc5, unlike other ORC subunits, when ectopically tethered to a chromatin locus, induces large-scale chromatin decondensation. The chromatin unfolding function of Orc5 requires its C-terminal domain but is independent of its AAA domain. Orc5 associates with the H3 histone acetyl transferase GCN5 and this association enhances the chromatin opening function of Orc5. In the absence of Orc5, histone H3 acetylation is decreased at the origins. I propose that Orc5’s ability to induce chromatin unfolding allows the establishment of the preRC at the origins.

In Chapter V, I summarize my findings on eukaryotic chromatin organization and DNA replication. In addition, I discuss several interesting avenues of explorations that these findings have opened up.
TABLE OF CONTENTS

CHAPTER I. INTRODUCTION .................................................................................................................. 1

CHAPTER II. UNDERSTANDING THE ROLE OF ORCA IN HETEROCHROMATIN ORGANIZATION AND REPLICATION - INTERACTION WITH H3K9 LYSINE METHYLTRANSFERASES .............................................................................................................................. 35

CHAPTER III. UNDERSTANDING THE ROLE OF ORCA IN HETEROCHROMATIN ORGANIZATION AND REPLICATION - INTERACTION WITH H4K20 LYSINE METHYLTRANSFERASES .............................................................................................................................. 99

CHAPTER IV. UNDERSTANDING THE ROLE OF Orc5 IN CHROMATIN DECONDENSATION ........................................................................................................................................................................... 114

CHAPTER V. CONCLUSIONS AND PERSPECTIVES .................................................................................. 140
CHAPTER I. INTRODUCTION

1.1. The Eukaryotic pre-Replication Complex

Accurate DNA replication is crucial for genomic integrity and cell survival. In eukaryotes, replication initiates at specific sites on the genome called origins of DNA replication. These sites are marked for replication initiation by a complex called the pre-Replication Complex (pre-RC) (Bell and Dutta, 2002; DePamphilis, 2005). The first step of replication initiation is the binding of a hetero-hexameric protein complex called Origin Recognition complex (ORC) to origins of DNA replication (Bell and Stillman, 1992). ORC then recruits Cdc6, a homolog of Orc1 that has an AAA+ motif (Neuwald et al., 1999). Cdc6 binds to ATP and this event is necessary for Cdc6’s association with chromatin (Herbig et al., 1999; Perkins and Diffley, 1998; Weinreich et al., 1999). In addition, hydrolysis of ATP is necessary for the downstream events of replication initiation. Cdc6 also increases the specificity of ORC by preventing its non-specific DNA binding and thereby playing a very important, albeit indirect role in origin selection. Cdc6 in turn recruits Cdt1. Cdt1 loads the replicative helicase Mini chromosome maintenance (MCM 2-7) complex onto origins (Maiorano et al., 2000; Nishitani et al., 2000) and at this stage origins are said to be licensed.

Origin selection in higher eukaryotes

ORC was first identified in yeast where its binding to origins is dictated by sequence specificity. In S.cerevisiae there are specific sequences called autonomously replicating sequences (ARS) (Newlon, 1988) which provide binding sites for ORC. ARS have four domains – A, B1, B2 and B3. DNase footprinting revealed the binding of ORC to A elements (Bell and Stillman, 1992). While sequence specificity is a feature of origins in S.cerevisiae, in higher eukaryotes what renders origins their identity is a subject of intense investigation and speculation. Another factor complicating the picture is that while ORC binds DNA in yeast, mammalian ORC does not have DNA binding.

Chapter 1.2 has been published with few modifications as:
capability. Chromatin, the structural and functional entity comprising DNA and histones, is gaining more prominence as a modulator of origin identity. Of gaining importance in the field of DNA replication initiation and origin selection are various factors which modify chromatin. To understand the possible roles that chromatin can play in influencing and dictating the process of replication, it is crucial to gain a deeper understanding of the canonical functions of these chromatin modifying machineries. So in the following section I will describe the major modifiers of chromatin organization and the mechanism by which they regulate crucial cellular processes.

1.2. Regulation of Chromatin organization and its role in DNA replication

Chromatin, the complex entity of DNA and protein, requires the packaging of DNA into a structurally organized and compact unit that enables the efficient progression of regulatory processes of the cell including transcription, DNA replication and repair. At its most elementary level chromatin consists of 146 bp of DNA wound 1.6 times around a histone core comprising two H2A-H2B dimers and a H3-H4 tetramer (Luger et al., 1997). As we examine the successive levels of complexity, chromatin transforms into a gigantic hub of protein-protein and protein-DNA interactions, thereby acquiring the ability to regulate myriad cellular processes in many different ways. For example, the ability to regulate proteins’ accessibility to DNA, to act as a scaffold for proteins to interact and its dynamics, all have major roles to play in replication, repair and transcription. Chromatin is distinguishable in two forms, euchromatin and heterochromatin. While euchromatin typically represents early replicating, transcriptionally competent and decondensed state of chromatin, the heterochromatin represents gene poor, late replicating and transcriptionally silent condensed chromatin. Here we focus on how chromatin modulates two fundamental processes, duplication of genetic material and control of gene expression.

Histone modifications and histone modifying enzymes:

Histone modifications play crucial roles in dictating the transcriptional status of a gene. Histones are subject to a variety of post translational modifications (PTMs) including methylation of lysines and arginines, acetylation of lysines, phosphorylation of serines and threonines, ubiquitylation and sumoylation of lysines, ADP ribosylation of glutamic
acid, deimination of arginines and isomerization of prolines (Kouzarides, 2007). These specific modifications especially at the histone tails could be associated with active or repressed transcription (Table 1). Typically, acetylation of H3 and H4 is associated with activation of transcription. The enzymes that catalyze acetylation (predominantly falling into the GNAT, MYST and CBP/p300 families) (for review, to (Roth et al., 2001)) can generally modify lysines at more than one position on the histones though there are examples of enzymes which do show some specificity. Unlike acetylation, methylation of lysines can activate or repress transcription depending on the residue that is modified (for review of histone modifications, refer to (Kouzarides, 2007)). Also, the enzymes, which catalyze methylation not only show exquisite specificity to particular lysine residues but also differ on the number of methyl groups they add (which can be mono, di and tri-methylation). Examples of activating methylation include those on H3K4, H3K36 and H3K79 (Bannister et al., 2005; Barski et al., 2007). The modified residue, as well as its location within the body of a gene plays crucial roles in modulating gene expression. For example, H3K4 trimethylation at promoter regions and H3K4 dimethylation and H3K36 and K79 trimethylation within the open reading frame (ORF) are associated with actively transcribing genes (Bannister et al., 2005; Barski et al., 2007). In contrast, the methylation marks associated with repression include H3K9, H3K27 and H4K20 (Barski et al., 2007). H3K9 di and tri methylation are involved in euchromatic gene silencing and heterochromatin formation, respectively, while di- and tri-methylation of H3K27 are mainly involved in transcription repression (Kouzarides, 2007; Schneider and Grosschedl, 2007). Though H4K20 trimethylation is generally thought to be involved in transcription repression, its function might actually be to maintain heterochromatic structures (Yang and Mizzen, 2009). While di- and tri-methylation of H3K9 and H3K27 are associated with repression, their monomethylated forms, when present within the ORF of a gene, could positively influence transcription (Barski et al., 2007).

This brings us to a concept, which is gaining credence in the field – that specific modifications on histones (example, methylation) do not necessarily “code” for a single kind of readout (example, repression) all of the time. Even though it is possible to divide histone modifications as those associated with active or repressed chromatin, the distinction is not always so clear (Schneider and Grosschedl, 2007). For example, H3K9
trimethylation and gamma isoform of HP1 have been shown to be associated with the ORFs of actively transcribing genes (Vakoc et al., 2005). Also, recent studies have demonstrated the role of HP1 in active transcription within euchromatin as well as for expression of genes that are harbored within heterochromatin (for review, (Kwon and Workman, 2011a; Kwon and Workman, 2011b)). There are also chromatin domains where both active and repressive marks are found. For example, genes that are not expressed in embryonic stem cells, but can be expressed in differentiated progenies show the activating marks of H3K4 methylation and H3K9 acetylation concomitant with the repressive mark of H3K27 trimethylation. The repressive mark may possibly be a part of the cellular mechanism to prevent the expression of specific gene(s) while the activating marks may be priming the same gene for expression in specific progeny (Azuara et al., 2006; Bernstein et al., 2005).

With so many modifications possible--not only on all histones but also on specific residues, --the importance of crosstalk between modifications arises. The presence of a particular modification on a residue (H3K9 acetylation) may preclude another modification (H3K9 methylation) from being created on the same residue. Modification on a residue can also hinder protein binding to an adjacent modified residue. This is the case when H3S10 phosphorylation prevents the binding of HP1α to an adjacent methylated H3K9 (Fischle et al., 2005). Modification of a particular residue can also positively and negatively affect the activity of an enzyme modifying a nearby residue (Clements et al., 2003; Nelson et al., 2006). Finally, not only does cross-talk occur within the same histone tail but also between tails, frequently referred to as “trans-tail”. One of the most well-known instances of trans-tail regulation is the regulation of H3K4 and K76 methylation by H2BK120 ubiquitylation (Laribee et al., 2007). More recently, H2BK120 Ub itself has been shown to be regulated along with H3K4 and K76 methylation by H2BK34 ubiquitylation (Wu et al., 2011).

Histone modifications can exert their effects through various mechanisms. Firstly, the modifications may have a direct effect on histone-DNA contacts both within a nucleosome and between nucleosomes by changing the net charge on nucleosomes. This can affect the structure at the nucleosomal level as well as the higher order chromatin
structure. For example, H4K16 acetylation has been shown to inhibit the formation of 30nm fibers (Shogren-Knaak et al., 2006). Histone modifications can also exert their influence by recruiting specific proteins that utilize specialized domains to recognize specifically modified histones. This in turn can alter chromatin structure and modulate gene expression, described as the “histone code hypothesis” (Jenuwein and Allis, 2001). The problem with investigating the role of histone modifications in transcription regulation is that it is difficult to prove whether a histone modification is the cause or consequence of transcription regulation. Also, the presence of redundant histone modifying enzymes and multiple histone genes in mammals makes the analysis of the role of histone modifications difficult. Apart from this many questions still remain, including the role of histone modifications in tethering genes to specific regions of the nucleus and the kind of modifications that are lost or gained when genes loop out of chromosomal territories during transcription (Schneider and Grosschedl, 2007). Future work will certainly enhance our appreciation of the crucial and diverse roles played by histone modifications in gene expression control.

**Chromatin remodeling factors:**

While histone modifying enzymes covalently modify histone tails, chromatin remodelers are proteins which can read these modifications and use the energy of ATP hydrolysis to change the interaction between histones and DNA, thereby regulating the access of the transcriptional apparatus to DNA (Clapier and Cairns, 2009). Remodelers fall into 4 classes including SWI/SNF, ISWI, NuRD/Mi-2/CHD and INO80 depending on the presence of additional domains in or near their ATPase domain (for review of chromatin remodelers refer to (Clapier and Cairns, 2009)). The remodeling activities may result in the eviction of histones, sliding of nucleosomes or complete removal of nucleosomes from specific DNA elements (Clapier and Cairns, 2009). This can regulate transcription by changing the accessibility of chromatin to transcription factors.

Eviction of nucleosomes is important for transcription initiation as supported by various reports of transcription factors binding to nucleosome-free regions (Mito et al., 2005; Yuan et al., 2005). This has been observed at the promoter of several specific gene loci such as the PHO5 promoter (Boeger et al., 2003; Reinke and Horz, 2003), as well as in
genome-wide studies in *S. cerevisiae* where promoters were found to be usually free of nucleosomes and flanked by positioned nucleosomes (Yuan et al., 2005). Nucleosome eviction is not only modulated by chromatin remodelers like Swi/Snf but also by several other factors including the sequence of DNA associated with the nucleosome (Drew and Travers, 1985; Stockdale et al., 2006). Co-activators like SAGA also play a role in nucleosome eviction at promoters (Govind et al., 2007). The HAT associated with the SAGA complex, GCN5 mediates eviction of nucleosomes and also increases H3K4 trimethylation within the gene body (Govind et al., 2007). Since the evicted histones can rebind to the same sequences that they were earlier associated with, histone chaperones (discussed in the subsequent subheading) like Asf1 and Nap1 are required for preventing futile nucleosome eviction (Adkins et al., 2004; Boeger et al., 2004; Lorch et al., 2006). Sliding of nucleosomes is also important for transcription regulation. For example, in yeast, ISW2 aids transcription repression by causing sliding of nucleosomes onto transcription start sites (Whitehouse and Tsukiyama, 2006). The role of ISW2 is particularly interesting in this case because it helps in sliding nucleosomes onto DNA sequences, which by themselves do not favor nucleosome assembly. Finally, eviction of histone is required for the movement of RNA polymerase on transcribing genes. This is supported by the observation that in *Saccharomyces cerevisiae* there is a negative correlation between histone density and the presence of RNA polymerase on transcribing genes (Schwabish and Struhl, 2004). But nucleosome free regions within a gene cannot remain so for a very long duration as it would lead to initiation of transcription from within the gene body. Accordingly, the eviction process is in equilibrium with deposition of histones at the wake of the transcribing RNA polymerase (Schwabish and Struhl, 2004). For this, chromatin remodelers have to work together with histone chaperones as discussed in the next section.

Not all cases of nucleosome remodeling favor active transcription. *In vivo* studies have shown that the ISW2 remodeling complex can move nucleosomes to positions that can block transcription (Whitehouse and Tsukiyama, 2006). *In vitro* data too points towards a role for these remodelers in limiting promoter activity (Stockdale et al., 2006; Zofall et al., 2006). Another example is that of NoRC, a member of the ISWI family of remodelers, which has a role in repressing rDNA gene transcription (Strohner et al.,
Apart from transcription activation and repression, remodelers also play key roles in transcription termination. In *S. cerevisiae*, transcription termination depends on the redundant actions of ISWI and CHD ATPases (Alen et al., 2002).

**Histone chaperones:**

Histone chaperones (HC) facilitate assembly, replacement or exchange of histones so as to organize the chromatin for transcriptional activation or repression (De Koning et al., 2007). Earlier it was thought that HCs were merely carriers of histones but in the past few years their roles in multiple aspects of gene regulation beginning to be discovered. It is now known that HCs are important for transcription initiation, elongation, prevention of non-specific transcription (within the body of a gene), transcription repression and heterochromatic spreading (Avvakumov et al., 2011).

For transcription initiation, HCs work in concert with chromatin remodelers to remove nucleosomes from promoters (Lorch et al., 2006; Nagaich et al., 2004). The HC, Asf1 aids both replication-coupled nucleosome assembly (discussed later in this review) and nucleosome eviction at promoters during transcriptional activation (Adkins et al., 2004; Boeger et al., 2004). This transcription aiding function of Asf1 has been attributed to its ability to bind to the exposed C-terminus of histone H4 after H2A-H2B dimers have been removed and then prying apart H3-H4 dimers (English et al., 2006).

Apart from this direct role in nucleosome eviction, Asf1 also influences transcription by regulating H3K56 acetylation (Williams et al., 2008). For many genes H3K56 acetylation has been shown to be crucial in transcription initiation (Rufiange et al., 2007; Schneider et al., 2006; Xu et al., 2005). For example, in yeast the absence of H3K56 acetylation leads to a reduction in recruitment of Snf5, a SWI/SNF subunit, to histone promoters and this in turn results in lower levels of H2A and H2B transcripts (Xu et al., 2005). The role of Asf1 in regulating H3K56 acetylation, influences induction of *PHO5* and *PHO8* genes (Adkins et al., 2004; Williams et al., 2008). This may well turn out to be a common mechanism for gene regulation. Another HC, FACT, has been shown regulate transcription elongation (Orphanides et al., 1998). FACT has been shown to remove H2A-H2B dimers from nucleosomes (Belotserkovskaya et al., 2003) and is also necessary for deposition of histones on DNA (Belotserkovskaya et al., 2003). FACT aids
transcription elongation by evicting H2A-H2B dimers from the path of RNA Polymerase II (RNAPII) and also prevents spurious transcription from within the gene body by restoring the original chromatin structure in the wake of RNAPII (for review on functions of FACT, (Formosa, 2011; Winkler and Luger, 2011).

HCs also regulate transcription elongation by modulating specific histone marks and function co-operatively with these marks. For example, FACT is necessary for monoubiquitination of H2B and this modification in turn is required for retaining FACT at the site of active transcription (Fleming et al., 2008). Another HC Spt6 is required for di- and tri-methylation of H3K36 (Carrozza et al., 2005). This modification in turn is recognized by the plant homeoboxdomain(PHD) and chromodomain of the HDAC complex Rpd3S and leads to its recruitment to that chromatin site (Li et al., 2007). This HDAC removes the activating acetylation marks from the transcribed chromatin and in turn leads to chromosome recompaction.

HCs are also required for transcriptional silencing and heterochromatin spreading. Asf1 associates with LID demethylase, which is responsible for removing H3K4 di- and trimethylation marks associated with active chromatin (Moshkin et al., 2009). Another HC, Nap1 binds to a complex containing the H3 histone deacetylase RPD3S (Moshkin et al., 2009). Asf1 and another HC, HIRA are also known to form a complex, which interacts with a HDAC Clr6 and leads to histone deacetylation and heterochromatin spreading (Yamane et al., 2011).

Understanding the detailed mechanism of how each of these HCs bind to their substrates, facilitate nucleosome assembly and transfer histones from one HC to another HC will provide important insights on the role of HC in modulating gene regulation.

**Histone variants:**

In addition to histone modifications, histone variants represent an important way to mark chromatin (for review,(Henikoff et al., 2004)). All histones except H4 have variants; examples include H1° and H5 for H1, H2A.X and H2A.Z for H2 and CenH3 and H3.3 for H3 (for review, refer to (Kamakaka and Biggins, 2005)). Many of the histone variants regulate transcription. For example, the H1 variant H5, which is found in chicken erythrocytes, is associated with the repression of transcription initiation *in vitro* and is
generally absent from active gene chromatin. Also, its deposition on chromatin during erythrocyte development coincides with transcription repression on a global scale (Wagner et al., 1977).

The H2 variant H2A.Z has been shown to have roles in both transcription activation and repression. It is usually deposited by the remodeler SWR1 near promoters in the nucleosomes that flank the nucleosome free regions (NFR) and might aid transcription by having a high turnover rate. The high turnover rate may expose the DNA sequences associated with these nucleosomes to various transcription factors thereby activating transcription (Raisner et al., 2005; Zhang et al., 2005). H2A.Z also has a role in transcription repression as suggested by the observations that its depletion leads to loss of Hp1α (Fan et al., 2004) and that mutations in H2A.Z disrupts Hp1α and polycomb protein mediated transcription repression (Swaminathan et al., 2005).

The variant of histone H3, H3.3 also has role in transcription regulation. In Drosophila and mammals it has been shown to be enriched at the promoter and ORF of actively transcribing genes (Chow et al., 2005; Goldberg et al., 2010; Mito et al., 2005; Schwartz and Ahmad, 2005) and also at the promoters of inactive genes, possibly marking them as poised for transcription (Mito et al., 2005; Tamura et al., 2009). These data support the idea of H3.3 having a role in transcription activation but in mouse ES cells, H3.3 is localized to telomeres and is required for the repression of telomeric repeats (Goldberg et al., 2010). So further investigations of the role of H3.3 at different chromatin domains would shed light on the different roles it may play in transcription regulation. Another histone H3 variant, Centromere protein A (CENP A) also influences epigenetic inheritance of the centromere in mammals (Palmer et al., 1991). CENP A’s counterpart CENH3 is found in all eukaryotes (Henikoff et al., 2004) and even though the satellite repeats constituting centromeres have rapidly evolved, CENH3 continues to be the identifying mark of centromeres including those which lack satellite repeats (example – human neocentromeres (Saffery et al., 2003)) indicating the power of this histone variant in propagating epigenetic information. Further work needs to be conducted in order to understand the various ways in which histone variants regulate transcription and the mechanisms by which they perform their roles as carriers of epigenetic information.
Chromatin and replication:

As with transcription, chromatin also affects various aspects of replication. The local chromatin landscape defines how origins of DNA replication are selected and activated so that duplication of the genetic material occurs accurately (Ding and MacAlpine, 2011). Also, during the process of replication, not only does an organism’s DNA have to be faithfully replicated but so does the chromatin structure and the epigenetic marks that are associated with it.

Chromatin context and origin selection:

Chromatin exerts its influence on replication right from the step of origin selection. The first evidence for this was the observation that of all the ARS (autonomously replicating sequences) in yeast, only a subset acted as active replication origins (Newlon et al., 1993). In metazoans the number of origins far exceed the origins that get activated, depending on various cellular needs including stress (Mechali, 2010). As with promoters, origins need to be nucleosome free for replication to occur. First, the origin recognition complex (ORC) is required for positioning the nucleosomes that flank origins while keeping the origins nucleosome free (Lipford and Bell, 2001). Secondly, origin sequences themselves encode information to keep the region nucleosome free (Eaton et al., 2010). Finally, transcription factors can function to keep promoters and origins free of nucleosomes. For example, yeast Abf1 has been shown to be required for creating nucleosome-free ARS and also for phasing of nucleosomes in the region flanking the ARS (Lipford and Bell, 2001). Not only do ORC binding sites have to be free of nucleosomes but there should also be phased nucleosomes flanking the binding sites. This phasing of nucleosomes is critical for formation of the pre-replication complex (preRC) even though it is not required for ORC binding per se (Lipford and Bell, 2001). Recent ChIP-seq studies in yeast have revealed that nucleosome-free regions and ORC-dependent phasing of nucleosomes are crucial factors for determining origins (Eaton et al., 2010).

What specifies an origin in higher eukaryotes (Figure 2) has been and continues to be an area of intense research (Cvetic and Walter, 2005). Unlike the yeast Saccharomyces cerevisiae, which has sequence-specific origins, no consensus sequence for origin
specification has been found in metazoans or even in *S. pombe* apart from the fact that origins are AT-rich. Recent reports have shown the role of specific histone modifications in establishing replication origins (Abbas et al., 2010; Centore et al., 2010; Tardat et al., 2010) (Table 1). The activity of H4K40 monomethylase PR-SET7 has been shown to be required during mitosis for setting the origins that are to be used in the next cell cycle. How PR-SET7 does this remains to be answered. One way could be that the mono-methylation mark created by PR-SET7 recruits the components of pre-RC by creating a suitable chromatin environment (Brustel et al., 2011). Another mark, which plays a role in origin licensing is H4 acetylation (on residues K5, 8 and 12) by the HAT HboI (Aggarwal and Calvi, 2004; Miotto and Struhl, 2010). This acetylation mark is found at origins with an H4K20 monomethylation mark during mitosis. Unlike PR-SET7 whose levels are cell cycle regulated (Rice et al., 2002) and whose over-expression causes DNA re-replication (Tardat et al., 2010), neither is the level of HboI cell-cycle regulated nor does its over expression cause re-replication (Iizuka et al., 2006; Miotto and Struhl, 2008). Together, these observations suggest that a sequence of histone modifications – H4K20 monomethylation during mitosis followed by H4 acetylation on several lysine residues could help in origin licensing (Brustel et al., 2011). After licensing, the degradation of PR-SET7 could prevent re-replication by not allowing the succession of histone H4 modifications to happen. Apart from H4K20 monomethylation, methylation on other histone residues also regulates origin selection (for review, (Dorn and Cook, 2011)) (table 1). For example, in budding yeast, H3K36 trimethylation is reduced at early firing origins (Pryde et al., 2009). In humans, it has been found that H3K4 trimethylation levels increase at early replicating origins (Birney et al., 2007; Karnani et al., 2010).

In addition to histone modifications, DNA methylation also influences origin selection. Origins, as well as promoters, are associated with CpG islands (Delgado et al., 1998). CpG islands may be involved in determining replication timing as origins having unmethylated CpG islands replicate faster compared to origins with methylated CpG islands (Gomez and Brockdorff, 2004).

There is also a connection between transcription and origin selection. Active transcription has been shown to repress replication while at the same time it has been shown that when origins are located at promoters, they fire easily, probably because of enhanced chromatin
accessibility (Ghosh et al., 2004; Kalejta et al., 1998). There are also various examples of transcription factors (TFs) being required for replication. For example, in D. melanogaster various TFs such as MYB, E2F and RB regulate DNA amplification at the Chorion locus (Beall et al., 2002; Bosco et al., 2001). Recently, work on identifying new origins by purifying origin centered nascent strands followed by hybridization to tiling arrays has again highlighted the interplay between replication and transcription (Karnani et al., 2010). In this study it was found that regions showing origin activity were near transcription start sites and were within 5 kb of transcription factors binding sites. Also origins were found to be enriched with marks usually associated with transcription initiation, namely, H3K4 di- and tri-methylation and H3 acetylation (Karnani et al., 2010). This becomes even more evident when we compare the set of common chromatin modifications associated with both transcription and with replication (table 1). Such an analysis strikingly shows that modifications associated with transcription activation are predominantly associated with early origins and have a positive correlation with replication timing.

Finally, a reason why no consensus for metazoan origins specification has been discovered might be because origin specification at various regions of the genome might have different requirements in terms of sequences and proteins involved (Mechali, 2010). Such a scenario would be similar to what happens during transcription in terms of different promoters recruiting different transcription factors. Such differences in origin selection could be due to several factors of which chromatin context could be an important one, as indicated by the growing body of work showing the influence of chromatin on replication.

**Inheritance of epigenetic marks: Replication coupled and independent chromatin assembly**

Once an origin fires and the replication machinery moves forward, the chromatin structure undergoes a massive change. Nucleosomes are removed from the path of the moving replication fork and in order for the and the parent and daughter DNA to receive their full complement of histones, newly synthesized histones need to be deposited along with the parental histones behind the replication fork (Probst et al., 2009). At this point,
the cell faces the challenge of passing on its epigenetic information to its progeny. For a mark to be truly epigenetic, it should be heritable. So various types of epigenetic marks such as DNA methylation and histone modifications must be faithfully copied from the parent to the daughter cells. But here is where a crucial difference between epigenetic and genetic information arises: while genetic information has to be faithfully transmitted, the replication of epigenetic marks can also provide an opportunity for these marks to be modified or erased, thereby changing the fate of the daughter cell (Probst et al., 2009). This flexibility of inheritance is very important for cells undergoing differentiation and also during various stages of development when a set of genes has to be turned off and a new set turned on. Thus, the process of DNA replication is crucial not only for DNA per se but also for regulated transmission of epigenetic marks because in the wake of the replication fork, epigenetic marks need to be restored or modified, as the situation may necessitate. The mechanism of restoration varies depending on the type of epigenetic mark. For example, DNA methylation on CpG islands is generally symmetrical on the parent strands and so the newly synthesized daughter strands are hemimethylated (Probst et al., 2009). The parent strand can be used as a template for DNA methylation of the daughter strand, that is, this epigenetic mark can be propagated by semi-conservative mode of information transfer similar to DNA replication. In mammals, the DNA methyltransferase DNMT1 is recruited to the sites of hemimethylated DNA by the SET and RING associated (SRA) protein NP45, which binds to hemimethylated DNA (Bostick et al., 2007; Sharif et al., 2007). In *A. thaliana*, recruitment of DNMT1 to the sites of hemi-methylated DNA is performed by the protein Variant in methylation1 (VIM1). Methylation by DNMT1 also requires the chromatin remodeler decreased DNA methylation 1 (DDM1) in *A. thaliana* and LSH in mice (Brzeski and Jerzmanowski, 2003; Dennis et al., 2001; Jeddeloh et al., 1999). The chromatin remodeling activity of DDM1 might be required for allowing the DNA methyltransferase to access the substrate. When considering the deposition of nucleosomes with their associated modifications at the replication fork, a problem arises. Since the parental nucleosomes are also disrupted during the passage of the replication fork, there is no obvious template for the deposition of appropriate nucleosomes. In the parental nucleosome, H3-H4 tetramers split from the H2A-H2B dimers (for review on nucleosome assembly, refer to (Franco and Kaufman,
In order to propagate epigenetic information accurately to the next generation, it is crucial that deposition of parental and newly synthesized histones is properly co-coordinated. Here histone chaperones (HC) play a crucial role. The HC CAF1 associates with H3 – H4 dimers and is required for their deposition on newly synthesized DNA (Smith and Stillman, 1989). The HC ASF1 associates only with newly synthesized H3-H4 dimers and serves as a donor to CAF1 (Mello et al., 2002). Parental H3-H4 tetramers or re-associated H3-H4 dimers can be deposited as such behind the replication fork, in which case the newly deposited nucleosome will have an old H3-H4 tetramer or, they can dissociate into dimers and associate with newly synthesized H3-H4 dimers thereby forming mixed nucleosomes. Apart from CAF1 and ASF1, the HC FACT could also play a crucial role in replication-coupled nucleosome assembly by associating with the DNA replication machinery component, the MCM helicase. It could also destabilize nucleosomes ahead of the replication fork and contribute to its eventual reassembly (review, (Formosa, 2011)). In yeast, a more recently described HC, Rtt106 has been proposed to have a role in both replication-coupled and replication-independent nucleosome assembly (Imbeault et al., 2008; Li et al., 2008). It functions similar to CAF1 in that it accepts H3-H4 dimers from ASF1 and deposits them behind the replication fork (Burgess and Zhang, 2010). This brings us to the next step: how are the post translational modifications (PTMs) of the parental histones copied onto the newly synthesized histones? Though an area of intense research, the principles governing nucleosome deposition and epigenetic inheritance in the wake of the replication fork are still far from clear. There are 2 models of how appropriate histone modifications are made on the newly synthesized histones deposited behind the replication fork (Margueron and Reinberg, 2010). The models are random and semiconservative. In the random distribution scenario, parental histone H3-H4 tetramers or dimers are distributed randomly on both the strands and they provide a template for modifying newly synthesized histone H3-H4 tetramers or dimers by inter- and intra-nucleosomal interactions, respectively. This model is not feasible because it would lead to gradual dilution of the post-translational marks (PTMs) of the parental histone. In the semiconservative mode, old and new dimers and tetramers of histones are shared equally by the DNA strands and there are two ways in which epigenetic information of the
histone PTMs may be transmitted. Firstly, old H3-H4 dimers might associate with newly synthesized H3-H4 dimers, forming mixed tetramers and the required PTMs can be made on the newly synthesized histones by using the corresponding old histone tail within the same nucleosome as template. In the second possible way of semi-conservative inheritance, old H3-H4 tetramers are deposited as such and are used by the new tetramers as template for modifications. Further work needs to be done to conclusively prove, which of the above-mentioned models reflects replication coupled nucleosome assembly in vivo.

While the canonical form of histone H3, H3.1/2 is deposited in a replication-coupled manner onto the chromatin, the variant H3.3’s deposition follows a replication-independent mechanism. Two histone chaperones, HIRA and Daxx play an important role in this process. HIRA was the first chaperone identified for the assembly of H3.3-H4 into nucleosomes (Ray-Gallet et al., 2002; Tagami et al., 2004). The other player, Death domain associated protein (Daxx) has also been identified to function as a HC for H3.3 (Drane et al., 2010; Goldberg et al., 2010; Lewis et al., 2010). HIRA and Daxx function in depositing H3.3 at different genomic locations. While HIRA is required for H3.3 incorporation at genic or transcribed regions, Daxx regulates the telomeric incorporation of H3.3 (Goldberg et al., 2010). Further experiments need to be done to shed light on the mechanism by which these histone chaperones deposit H3.3-H4 at specific genomic locations. It would also be interesting to investigate the possible crosstalk and competition between the various H3.3-H4 HC (CAF1, HIRA and Daxx) complexes and how they might regulate each other’s functions.

Influence of replication proteins on Chromatin: heterochromatin formation

Not only does chromatin exert its influence on replication but replication proteins also affect chromatin architecture (for review, (Chakraborty et al., 2011a; Probst et al., 2009)). ORC has a role in transcriptional silencing in *S. cerevisiae* (Dillin and Rine, 1997) and heterochromatin formation in several organisms (Auth et al., 2006; Ehrenhofer-Murray et al., 1995; Gerbi et al., 2002; Pak et al., 1997; Prasanth et al., 2004; Prasanth et al., 2010) (for review, (Chakraborty et al., 2011a)). In *S. cerevisiae* Orc1 recruits Sir1 to mating type loci, HML (Hidden MAT Left) and HMR (Hidden MAT Right) and is also required
for their transcriptional silencing (Triolo and Sternglanz, 1996). Mutation of the N-terminus of Orc1, which is required for recruiting Sir1, leads to loss of transcriptional silencing (Bell et al., 1995). The replication and silencing functions of ORC are separable in yeast (Dillin and Rine, 1997; Fox et al., 1997) though the silencing by ORC requires cells to pass through S phase (Dillin and Rine, 1997; Fox et al., 1997; Miller and Nasmyth, 1984). Mutation of Orc2 in S. cerevisiae causes disruption of silencing at HMR-E locus (Bell et al., 1993; Foss et al., 1993; Micklem et al., 1993). In Drosophila, mutation of Orc2 leads to inhibition of position effect variegation (Pak et al., 1997) and both in Drosophila and Xenopus, ORC subunits interact with Heterochromatin Protein1 (HP1) (Huang et al., 1998; Pak et al., 1997). In humans, various subunits of the ORC (Orc1, 2, 3 and 5) have also been shown to be localized to heterochromatic regions and interact with HP1 (Prasanth et al., 2010). Also, knockdown of several ORC subunits causes changes in HP1α distribution. That Orc2 and Orc3 are required for the association of HP1 with heterochromatin is supported by the observation that knockdown of these proteins leads to a homogenous distribution of HP1α in the cell. Conversely, HP1 knockdown in human cells leads to loss of Orc2 from heterochromatin indicating that HP1 and Orc2 (in association with Orc3) stabilize each other on heterochromatin (Prasanth et al., 2010). Another aspect of chromatin organization that ORC affects is the compaction of satellite repeats. It has been shown that knockdown of Orc2 and Orc3 leads to loss of compaction of the satellite repeats of chromosome 9 (Prasanth et al., 2010). On a separate note, in Drosophila, Orc2 has also been shown to be required for replication of euchromatic regions at their normal replication time during S phase. Orc2 mutants show delayed replication of euchromatin followed by aberrant condensation of chromosomes during mitosis (Loupart et al., 2000). This is a very interesting observation, as it links replication timing to chromatin structure.

More recently, an ORC-associating protein – leucine rich repeats and WD40 repeat domain-containing protein 1 alias ORC-associated (LRWD1/ORCA) has been identified (Bartke et al., 2010; Shen et al., 2010; Vermeulen et al., 2010), which has also been shown to co-localize with ORC at heterochromatin. ORCA interacts with ORC and knockdown of ORCA leads to reduced ORC loading onto chromatin (Bartke et al., 2010; Shen et al., 2010). ORCA and ORC have been shown to bind to the repressive H3K9,
H3K27 and H4K20 tri-methylation marks cooperatively with DNA methylation marks (Bartke et al., 2010; Vermeulen et al., 2010) and it is has been postulated that ORCA might mediate the interaction of ORC with heterochromatin by directly binding to repressive marks (Vermeulen et al., 2010). It would be interesting to investigate whether ORCA assisted ORC loading is the universal mechanism of ORC association with chromatin or whether it is a feature of specific chromatin subtypes like heterochromatin or with specific subsets of origins of replication.

There is also an enormous amount of crosstalk between replication and chromatin modifying machineries. For example, CAF1’s activity of depositing H3-H4 particles requires its association with the DNA polymerase sliding clamp, proliferating cell nuclear antigen (PCNA) (Krawitz et al., 2002; Shibahara and Stillman, 1999; Zhang et al., 2000). CAF1’s role in replication coupled nucleosome assembly is also supported by its S phase association with replication foci and with foci of PCNA after DNA damage (Martini et al., 1998; Shibahara and Stillman, 1999; Taddei et al., 1999). In human cells, CAF1’s depletion leads to DNA replication defects and the arrest of cells in S phase. These data together indicate that CAF1 provides the link between chromatin assembly and DNA replication (Hoek and Stillman, 2003; Ye et al., 2003). PCNA also plays a pivotal role in recruiting histone deacetylases, histone methyltransferase PR-SET7 (Huen et al., 2008; Jorgensen et al., 2007; Milutinovic et al., 2002), and chromatin remodelers to replication sites. For example, PCNA recruits Williams Syndrome Transcription Factor (WSTF) to sites of replication and WSTF in turn recruits SNF2H, a chromatin remodeler of the ISWI family (Poot et al., 2004). Also at regions where the daughter strand is hemi-methylated, PCNA along with NP95 recruits DNMT1 (Bostick et al., 2007; Chuang et al., 1997; Sharif et al., 2007). This in turn expands the network of interacting proteins at the replication fork, as DNMT1 can interact with G9a, a lysine methyl transferase (Esteve et al., 2006) that is required for H3K9 methylation. As more instances of overlapping functions of replication and chromatin modifying machineries are reported, a clearer picture of the regulation involved in these processes is emerging and more work is required to fully understand these highly interconnected events.
Replication and Transcription: Crosstalk on chromatin

Apart from the influence of transcription on origin selection discussed earlier, there is evidence from several quarters that the set cellular signals, or more accurately, chromatin signals, controlling transcription and replication may be one and the same. For example, in *Drosophila*, different studies point toward a link between replication timing and transcriptional status of the single X chromosome in males (Chatterjee and Mukherjee, 1977; Gorman and Baker, 1994; Schwaiger et al., 2009). In males, transcription of the X chromosome is increased so that the gene products of the single X chromosomes can be comparable in abundance to those transcribed from autosomes (Gorman and Baker, 1994). The histone acetyl transferase Mof aids this process by hyperacetylating the X chromosome to create a conducive environment for increasing the rate of transcription (Akhtar and Becker, 2000). Interestingly, it has been observed that this Mof-dependent hyperacetylation of H4K16 also alters the replication timing of the X chromosome, leading to the earlier replication of the chromosome compared to autosomes (Schwaiger et al., 2009). Similar results of early replication of the single X-chromosome in males have also been obtained through genome wide studies in *Drosophila* (MacAlpine et al., 2004; Schubeler et al., 2002; Schwaiger et al., 2009), mouse (Farkash-Amar et al., 2008; Hiratani et al., 2010) and humans (Ryba et al., 2010), thereby reiterating the intimate connection between replication and transcription where the chromatin context favoring transcription also favors replication.

Linking replication and transcription seems to be an efficient strategy of the cell. Transcription generates negative supercoiled DNA upstream of the gene being transcribed that could facilitate ORC binding (Ding and MacAlpine, 2011). This would be similar to what happens *in vitro* where it has been shown that ORC preferentially binds to negatively supercoiled DNA (Remus et al., 2004).

Though many new examples of the interplay between transcription and replication are being reported with increasing frequency, several important questions remain to be addressed about the exact nature of the cross-talk. How much of the “cross-talk” is actually an influence of transcription on replication versus the mere effects of creation of a suitable chromatin environment by the process of transcription needs to be determined.
1.3. ORCA/LRWD1: Connecting DNA replication to chromatin organization

A few years back our lab identified a novel ORC interactor called Origin Recognition Complex Associated (ORCA)/LRWD1. Previous work from our lab showed that ORCA interacts robustly via its WD domain with ORC (Shen et al., 2010). Levels of ORCA peak in G1 phase of cell cycle and decrease at G1/S boundary. Interestingly, the remaining ORCA after G1 associates with heterochromatin. This association with bona fide heterochromatic structures like telomeres and centromeres requires ORCA’s WD domain. ORCA associates with centromeres in telomerase positive cells (eg. MCF7, Hela and IMR-90). Interestingly, ORCA associates with telomeres in the interphase of cells that use Alternative Lengthening of telomeres (ALT) (eg. U2OS). In many kinds of cancers, cells maintain their ability to replicate and divide indefinitely by maintaining the levels of telomerase (Hanahan and Weinberg, 2011), a step that prevents telomere erosion. In some cancers, primarily sarcomas, cells use another mechanism involving homologous recombination termed as ALT (Bryan et al., 1997). This is very interesting because it could point towards a role of ORCA in DNA recombination or repair.

Supporting the hypothesis is published data showing ORCA is a substrate of ATM and ATR kinases (Matsuoka et al., 2007). During mitosis, ORCA associates with centromeres in both the kinds of cells (Telomerase+ and ALT+ cells).

Tethering of ORCA to an artificial locus leads to the recruitment of ORCs and MCMs indicating that de novo origins could be established upon tethering ORCA. In addition, loss of ORCA leads to loss of ORCs and MCMs from chromatin in both cancerous U2OS cells and primary diploid fibroblasts, WI38. Further analysis of cell cycle progression upon ORCA knockdown revealed that ORCA was required for entry into cell cycle.

ORCA interacts with multiple ORC subunits directly (Shen et al., 2012). In addition, ORCA also interacts with the pre-RC component Cdt1 and its inhibitor geminin. By using Single-Molecule Pull-down, it was demonstrated that one molecule each of Orc1 and Cdt1 interact with ORCA while two molecules of Geminin interact with ORCA. This data is in line with the known stoichiometry of pre-RC, adding to the line of evidence that ORCA possibly functions in replication initiation as a part of pre-RC. Another interesting piece of data is that in a cell, the level of Orc2 is several folds higher than the level of ORCA indicating the existence of ORCA independent ORC complexes. As a natural
segue, it would be extremely interesting to determine the specific role of ORCA containing ORC complexes. These complexes could be marking specific origins, for example, those residing in late replicating heterochromatin. Another, not mutually exclusive possibility is that ORCA-ORC complexes could be marking repressed chromatin environments in general. ORCA interacts with Orc1 and Cdt1 in G1, with Orc2 throughout cell cycle and with Geminin in post-G1 cells. In addition, ORCA interacts with phosphorylated Cdt1 in mitosis. These differential interactions of ORCA with the various pre-RC components could provide an additional level of regulation to the process of replication initiation and licensing. At G1/S, when the levels of Geminin increase, the interaction between ORCA and Cdt1 is lost and ORCA changes its interacting partner to Geminin. This could lead to the loss of Cdt1 from origins and could possibly be a mechanism by which cells prevent re-replication.

Interestingly, Orc2 stabilizes ORCA. Knocking down Orc2 leads to a decrease in ORCA levels. So to understand the mechanism of this regulation, previous work from our lab investigated the regulation of ORCA by the proteosomal machinery. As levels of ORCA decrease dramatically at G1/S, it could serve as a possible regulation module for replication initiation. Studies from our lab showed that ORCA is ubiquitinated at G1/S (Shen and Prasanth, 2012) via K48 linkage – a modification associated with proteosomal degradation. Orc2 binds to ORCA and protects it from this degradation step. Poly-ubiquitination of ORCA occurs on chromatin and ubiquitinated ORCA remains associated with detergent resistant structures in the nucleus. In addition, ORCA associates with the E3 ligase, Cul4A – Ddb1 and it is possible that this could be the E3 ligase mediating the ubiquitination of ORCA.

ORCA has the potential to turn into a major player linking DNA replication initiation and chromatin organization. It associates with members of the pre-RC and also with heterochromatin in post-G1 cells. In addition, it has a WD domain that can bind repressive histone modifications. In the following chapter, I will investigate the role of ORCA in heterochromatin organization and replication by exploring its interaction with the repressive H3K9 lysine methyltransferases.
1.4. Tables and figures

Table 1.1. List of modifications common to transcription and replication

<table>
<thead>
<tr>
<th>Mark</th>
<th>Function: measured as activation or repression of transcription</th>
<th>Mark</th>
<th>Function: measured as association with early/mid/late firing origins or correlation with replication timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3K4me1</td>
<td>Repression and activation</td>
<td>H3K4me1</td>
<td>Enriched at early origins</td>
</tr>
<tr>
<td>H3K4me2/3</td>
<td>Activation</td>
<td>H3K4me3</td>
<td>Enriched at early origins and depleted at late origins; may aid origin firing</td>
</tr>
<tr>
<td>H3K9me2/3</td>
<td>Repression and activation</td>
<td>H3K9me2/3</td>
<td>Negatively correlates with replication timing, that is, with early origins</td>
</tr>
<tr>
<td>H3K27me3</td>
<td>Repression</td>
<td>H3K27me3</td>
<td>Negatively correlates with replication timing, that is, with early origins</td>
</tr>
<tr>
<td>H3K26me3</td>
<td>Activation</td>
<td>H3K26me3</td>
<td>Increases at origins during S phase and aids origin firing</td>
</tr>
<tr>
<td>H4K20me1/2/3</td>
<td>Repression</td>
<td>H4K20me1</td>
<td>Inhibits origin firing; Enriched at late firing origins and depleted at early origins</td>
</tr>
<tr>
<td>H2BK23Ub</td>
<td>Activation</td>
<td>H2BK23Ub</td>
<td>High during mitosis and G1. Promotes origin licensing possibly via aiding acetylation of H4K5, B and D2</td>
</tr>
<tr>
<td>H3K9Ac</td>
<td>Activation</td>
<td>H3K9Ac</td>
<td>High levels at active origins</td>
</tr>
<tr>
<td>H4K16Ac</td>
<td>Activation</td>
<td>H4K16Ac</td>
<td>Required for activation of subset of origins</td>
</tr>
<tr>
<td>CpG islands</td>
<td>Activation</td>
<td>CpG islands</td>
<td>Early origins</td>
</tr>
</tbody>
</table>
Figure I.1. Determinants of origin specification – several factors influence the selection of origins of replication. These include the presence of AT rich sequences, nucleosome dynamics and the presence of phased nucleosomes flanking origins. ORC binds to origins and influences the phasing of nucleosomes. Transcription factors can recruit ORC and similar recruitment of ORC could be carried out by unknown players and the recently identified ORC-associated (ORCA/LRWD1). There is a growing body of data pointing towards the role of DNA methylation and histone modifications in origin specification.
1.5. References


Formosa, T. 2011. The role of FACT in making and breaking nucleosomes. *Biochimica et biophysica acta*.


Williams, S.K., D. Truong, and J.K. Tyler. 2008. Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional...


CHAPTER II. UNDERSTANDING THE ROLE OF ORCA IN HETEROCHROMATIN ORGANIZATION AND REPLICATION – INTERACTION WITH H3K9 LYSINE METHYLTRANSFERASES

II.1. Introduction

Origin Recognition Complex-Associated (ORCA/LRWD1), a protein required for the initiation of DNA replication, preferentially localizes to heterochromatic regions in post replicated cells (Bartke et al., 2010; Shen et al., 2010; Vermeulen et al., 2010). We and others have demonstrated that ORCA and ORC associate with centromeric and telomeric heterochromatin in mammalian cells (Shen et al., 2010). Further, using a stable isotope labeling by amino acids in cell culture (SILAC)-based proteomic approach, ORCA-ORC complex has been shown to bind the repressive histone lysine methylation marks, specifically H3K9me3, H3K27me3 and H4K20me3 (Bartke et al., 2010; Vermeulen et al., 2010) that are known to be enriched at heterochromatic sites. ORCA contains a WD-domain, a structure known to interact with nucleosomes/histones (Wysocka et al., 2005). We have previously demonstrated that the WD-domain of ORCA is crucial for its binding to heterochromatin. Furthermore, ORCA is critical for stabilizing ORC binding to chromatin (Shen et al., 2010). ORC, a hetero-hexameric complex, in addition to serving as the landing pad for the assembly of pre-replicative complex at the origins of DNA replication, participates in sister chromatid cohesion, heterochromatin organization, and chromosome segregation (Bell et al., 1993; Sasaki and Gilbert, 2007; Shimada et al., 2002). In metazoans, ORC also facilitates the association of heterochromatin protein 1 (HP1) to the H3K9me3-containing pericentric heterochromatin (Pak et al., 1997; Prasanth et al., 2004; Prasanth et al., 2010). Thus, it is obvious that ORC-ORCA complex associates with heterochromatin, but the mechanism underlying the recruitment of this multiprotein complex to the condensed chromatin and the functional relevance of such association has remained elusive for decades.

Histone lysine methylation, catalyzed by lysine methyltransferases (KMTs), plays key roles in the epigenetic regulation of chromatin organization, transcription and replication.

Chapter II has been published with few modifications as: Giri SG, Aggarwal V, Pontis J, Shen Z, Chakraborty A, Khan A, et al. The preRC protein ORCA organizes heterochromatin by assembling histone H3 lysine 9 methyltransferases on chromatin. eLife 2015; 4.
Methylation of H3K9 is an abundant and stable modification and is an important regulator of heterochromatin formation, gene silencing and DNA methylation (Martin and Zhang, 2005). The methyl modifications on H3K9 exist in distinct mono-, di- and tri-methyl states (H3K9me1, H3K9me2 and H3K9me3, respectively), with each responsible for governing distinct cellular functions. In general, H3K9me1 and H3K9me2 are associated with gene expression/repression at euchromatic regions, whereas the H3K9me3, enriched at pericentric heterochromatin, is required for heterochromatin assembly and gene silencing (Martin and Zhang, 2005). The major KMTs catalyzing these modifications are G9a and GLP, responsible for H3K9me2 (Rice et al., 2003; Shinkai and Tachibana, 2011; Tachibana et al., 2001); SETDB1, which establishes H3K9 di and trimethylation in euchromatin (Schultz et al., 2002) and Suv39H1/H2 that establish H3K9me3 from mono- or di-methylated H3K9 (Peters et al., 2003; Rea et al., 2000). While the idea of G9a and Suv39H1 acting in distinct, primarily in non-overlapping chromatin contexts held sway for a long time, this concept has been recently challenged by the discovery of a complex consisting of multiple H3K9 KMTs (Fritsch et al., 2010). The multimeric complex contains all four H3K9 KMTs G9a, GLP, Suv39H1 and SETDB1 and is recruited to both pericentromeric heterochromatin and promoter of a set of G9a repressed genes where it aids in gene repression by maintaining H3K9me2 and H3K9me3 marks (Fritsch et al., 2010). Furthermore, destabilizing even one of these KMTs resulted in the disintegration of the multimeric complex and loss of the enzymatic activity of this complex (Fritsch et al., 2010). How this multimeric KMTs complex is recruited to specific chromatin sites remained to be determined. In the broader context, the functional significance of the crosstalk between chromatin modifying and replication machineries has remained largely unexplored.

Here we demonstrate that ORCA associates with H3K9 KMTs in a chromatin context-dependent manner. By using a highly sensitive and quantitative Single-Molecule Pull-down (SiMPull) approach (Jain et al., 2011; Shen et al., 2012) we demonstrate that ORCA preferentially binds to H3K9me3 and ORCA-ORC and multiple H3K9 KMTs exist in a single complex. Furthermore, ORCA is required for the formation and/or maintenance of the H3K9 KMT complex. Our results indicate that ORCA is required for the integrity of global chromatin architecture. In the absence of ORCA, human cells show
alterations in the binding and activity of KMTs at sites enriched for these factors with concomitant reduction in H3K9me2 and H3K9me3 marks. Finally, we observe that the cells lacking ORCA display abnormal heterochromatin organization and alteration in the replication timing, specifically at the late-replicating regions. We propose that ORCA is a scaffold protein that is required for the establishment as well as maintenance of heterochromatin.

II.2. Results

II.2.1. ORCA interacts with H3K9 KMTs

In order to address if ORCA interacts with the machinery that causes the establishment of heterochromatin, we used a candidate approach to investigate the interaction of ORCA with individual H3K9 KMTs that catalyze H3K9 repressive modifications. We observed robust interaction of endogenous ORCA with endogenous G9a and Suv39H1 (Fig. 1Aa and b, Figure 2Aa and Ab). 1.31% of total G9a was found to be in a complex with ORCA. Quantitation was based on the amount of G9a immunoprecipitated with ORCA (based on 100% efficiency of ORCA IP, Figure 1- Figure Supplement 1B) (n=7). Similarly, 1.44% of total Suv39H1 was in a complex with ORCA (n=4). Note that only about 0.2% of the endogenous H3K9 KMTs co-purified with Suv39H1 (Fritsch et al., 2010). Co-immunoprecipitation (co-IP) using T7-tagged ORCA and Flag-tagged H3K9 KMTs revealed interaction of ORCA with G9a, GLP and Suv39H1, all enzymes involved in the establishment of heterochromatin (Fig. 1Ba and b). In addition, we carried out IP from cell lines stably expressing Flag-tagged-G9a or GLP. IP from nuclear extracts using Flag antibody to determine the association of endogenous ORCA with the KMTs. ORCA along with Orc2 and MCMs were found to interact with the KMTs (Figure 2C). However, ORCA did not associate with the arginine methyltransferase PRMT5 (Figure 2D), showing the specificity of the interactions.

In order to show functional co-recruitment of ORCA and the H3K9 KMTs, we used an *in vivo* cell system (CLTon) that uses a 200 copy transgene array-containing Lac operator repeats stably integrated into human osteosarcoma (U2OS) cells as a single heterochromatic locus that can be visualized by Cherry-Lac repressor (LacI). Further, transcriptional activation using doxycycline (DOX) causes the decondensation of the
locus (Janicki et al., 2004; Shen et al., 2010). We tethered the triple fusion proteins of YFP-LacI-KMTs to the heterochromatic locus and examined if these enzymes could recruit ORCA to the locus. This approach corroborated the interaction of ORCA with G9a (Fig. 1C).

We next examined whether ORCA and H3K9 KMTs (G9a and Suv39H1) assembly requires intact DNA. Co-IPs from cells expressing G9a and ORCA or Suv39H1 and ORCA were carried out in the presence or absence of Ethidium bromide (EtBr). EtBr selectively inhibits DNA-dependent protein interactions (Lai and Herr, 1992). ORCA continued to show interaction with G9a as well as Suv39H1 even in the presence of EtBr (Fig. 1Da and Db), indicating that these interactions were DNA-independent. The DNA independent interactions were also corroborated by co-IP experiments in the presence of the nuclease benzonase (data not shown). Further, the interaction of ORCA with G9a as well as SuV39H1 was direct and independent of DNA, as evident by the direct interaction of purified ORCA with G9a/SUV39H1 proteins (Fig 1Ea and Eb).

II.2.2. The association of ORCA and H3K9 KMTs occurs on condensed chromatin

Recent studies have demonstrated that in addition to Suv39H1, G9a/GLP may participate in the establishment of pericentric heterochromatin (Dong et al., 2008; Kondo et al., 2008; Vassen et al., 2006). Since ORCA is enriched at heterochromatic regions, we carried out detailed functional characterization of the interaction of ORCA with Suv39H1 and also with G9a in order to dissect the biological relevance of these associations.

To map the interaction domains of ORCA with G9a and Suv39H1 we generated several truncation mutants of ORCA (Fig. 3Aa), G9a (Fig. 3Ab) and Suv39H1 (Fig. 3Ac). Using co-IP experiments, we observe that the WD repeats of ORCA (truncation mutants 128-647 and 270-647 aa) interacted with G9a (Figure 4Aa) and Suv39H1 (Figure 4Ab). We found that the deletion of any one of the WD domains in ORCA resulted in loss of binding to heterochromatin consistent with the fact that the intact β-propeller structure of WD is crucial to maintain its functionality. We also observed that the Leucine-Rich Repeats (LRR)-containing fragment of ORCA (1-127aa), but not the one containing the linker (1-270 aa), interacted with G9a (Figure 4Aa) but not with Suv39H1 (4Ab). Co-IP experiments demonstrated that the Ankyrin repeat (619-965 aa) of G9a (Fig. 3Ba) and the
SET domain (151-412 aa) of Suv39H1 were necessary for interaction with ORCA (Fig. 3Bb).

To address if the chromatin context affected the interaction between ORCA and G9a we used the CLTon cells and examined the interaction of YFP-LacI-fused full-length and various truncation mutants of ORCA tethered to the heterochromatic locus, with full-length or truncation mutants of G9a (Fig. 3Ca-b, Da-b, Figure 4Ba-b). The WD domain of ORCA was able to recruit CFP-G9a (Fig. 2Ca-b) corroborating our IP results (Figure 4Aa). Interestingly, when CFP-LacI-ORCA was co-transfected with YFP-G9a mutants (Figure 4Ba), not only did the mutant YFP-G9a-1-618, which lacks the Ankyrin repeats, show significantly reduced interaction but also YFP-G9a-1-965, which has an intact Ankyrin repeat but lacks the SET domain, showed significantly reduced association with the locus (Figure 4Ba-b). We next addressed if the interaction of ORCA with G9a at heterochromatic regions required the catalytic domain of G9a in addition to its Ankyrin repeats. YFP-LacI-G9a triple fusion protein was found to be enzymatically active as is evident by the accumulation of H3K9me2 at the CLTon locus upon the tethering of G9a full-length construct (Figure 4C). Tethering of YFP-LacI-G9a-ΔSET or YFP-LacI-G9a-H1166K (a point mutant, which abolishes the catalytic ability of G9a) to the locus (Fig. 3Da), failed to recruit CFP-ORCA (Fig 3Da-b). However, co-immunoprecipitation experiments demonstrated that T7-ORCA could interact with GFP-G9a-ΔSET (Figure 4D). These data suggest that while the Ankyrin repeat of G9a is sufficient for the association with ORCA (Fig. 3Ba), the interaction of ORCA and G9a at the heterochromatin also requires the methylating ability of G9a. Similarly, the interaction between ORCA and Suv39H1 requires the SET catalytic domain (Fig. 3Bb).

Since we observed the interaction between ORCA and the H3K9 KMTs at the heterochromatic locus, we next asked if the interaction occurred in a chromatin context-dependent manner. We tethered ORCA to the CLTon locus and examined the recruitment of G9a upon induction of transcription, from the decondensed locus (Fig. 3Ea). In the absence of doxycycline, ~80% of cells showed CFP-G9a recruitment to the locus when YFP-LacI-ORCA was tethered (Fig. 3Ea-b). Upon transcriptional activation there was a striking reduction (~10%) in CFP-G9a association in the YFP-LacI-ORCA-tethered
decondensed locus (Fig. 3Ea-b). ORCA-tethered decondensed locus, in addition to G9a, also failed to recruit HP1α, but contained Cdk9, a component of the pTEFB kinase complex, which is part of the transcription elongation complex (Figure 4E). Two components of the ORC, Orc2 and Orc3 that require each other for their stability associate with one another at both the condensed and the open chromatin (Figure 4F). These results indicate that while ORCA can interact directly with G9a and Suv39H1 (Fig. 1Ea-Eb), the interactions could be dependent on or regulated by chromatin within the cells. In a cellular milieu, the fraction of ORCA directly interacting with G9a or Suv39H1, independent of chromatin is likely to be a very small pool, and therefore too weak to be detected at the CLTon locus.

II.2.3. ORCA-ORC and the H3K9 KMTs exist in one single complex

Our earlier work demonstrated the existence of a subset of multiple H3K9 KMTs in a single complex and functional cooperation between these molecules to regulate heterochromatin function and gene expression (Fritsch et al., 2010). Since ORCA interacts with different H3K9 KMTs, we investigated if ORCA is an integral component of this multi-KMT complex.

For this purpose we employed the process of Single molecule pull-down (SiMPull) analysis (Fig. 5Aa-b) (Jain et al., 2011). This method is extremely sensitive and is a tour-de-force to examine protein complexes and also to accurately calculate the stoichiometry of proteins within the complexes (Shen et al., 2012). This approach obviates the need for ensemble experiments that require IPs with large quantities of cell lysates. Our initial estimates predicted that three grams of Flag-HA-Suv39H1-expressing Hela-S3 cell pellet, which is ~3 billion cells, is required for a single glycerol gradient sedimentation to obtain other H3K9 KMT signals detectable by western blotting (Fritsch et al., 2010). However, a relatively higher sensitivity can be achieved by the SiMPull approach by using only a million cells. We first measured the stoichiometry of ORCA bound ORC and H3K9 KMTs, respectively. We used cells co-expressing T7-ORCA and YFP-ORC1 (Fig. 5Ba-d) or T7-ORCA and YFP-KMTs to perform SiMPull (Fig. 5Ca-d; 3Da-d). ORCA complexes containing YFP-ORC1 or YFP-KMTs were visualized as isolated fluorescent spots by single molecule Total Internal Reflection Fluorescence (TIRF) microscopy.
Pull down by a control antibody (anti-HA) showed very low non-specific level of fluorescence, thereby demonstrating the high specificity of SiMPull assay. Individual fluorescence spots showed single or multi-step decreases in fluorescence intensity corresponding to photobleaching of individual YFP molecules in a single complex (a representative schematic of the photobleaching analysis is shown in Fig. 5Aa). After photobleaching analysis of many co-immunoprecipitated YFP-ORC1, YFP-G9a and YFP-Suv39H1, we found that primarily one molecule each of ORC1, G9a and Suv39H1 interacts with a single molecule of ORCA (Fig. 5Bd, Cd, Dd). In addition, we also observed that in a small population, ORCA associates with two molecules of G9a (Fig. 5Cd and Ce), suggesting that the ORCA-interacting-G9a may also be present as a homodimer.

Next, we investigated whether ORCA bound to KMTs also contains ORC by performing SiMPull using biotin-conjugated anti-T7 antibody with cells co-transfected with YFP-ORC1, mCherry-G9a and T7-ORCA (Fig. 5Ea). To mimic endogenous expression levels of these proteins, we first systematically titrated the levels of plasmid transfected to obtain an expression of the candidate protein that is similar to endogenous levels (Figure 6Aa). Based on this analysis, we transfected 2X10^6 cells with 100ng of each plasmid and then carried out SiMPull (Figure 6Aa, lane3). Complexes of T7-ORCA that contain YFP-Orc1 were detected in the green imaging channel and those containing mCherry-G9a were detected in the red imaging channel (Fig. 5Eb-c). After overlaying the two channels, 39±5% of YFP-ORC1 molecules colocalized with mCherry-G9a, indicating that all three proteins, ORC, ORCA and G9a are found in a substantial fraction of single complexes (Fig. 5Ec). SimPull from cells that were transfected with higher concentration of plasmid showed similar extent of co-localization (Figure 6Ab-Ac). This results is consistent with our earlier study showing that ORCA is protected from ubiquitin-mediated proteolysis when bound to ORC and as a result is always associated with ORC (Shen and Prasanth, 2012). Next, we tested whether multiple H3K9 KMTs exist in a single complex with ORCA using triple transfections of T7-ORCA, YFP-Suv39H1 and mCherry-G9a in U2OS cells (Fig. 5Fa-c, plasmid titrations: Figure 6Ba, lane3 used for the experiment). Interestingly, we could observe ~55±7% of YFP-Suv39H1 colocalized with mCherry-G9a, (Fig. 5Fc). Similar results were obtained with higher concentration
of plasmid transfection: Figure 6Bb-Bc), suggesting the existence of a significant amount of ORCA-G9a-Suv39H1 complex. The true degree of cohabitation may be even higher because the fluorescent proteins may not all mature into active chromophores. This leads to dark molecules and appearance of either only green or only red spots even though both the KMTs are present in a complex. In addition, unequal expression of the transfected KMTs or the presence of endogenous KMTs in the complexes may also lead to a reduction in cohabitation detected by SiMPull. Finally, only a subset of G9a and Suv39H1 may exist as a single complex with ORCA [similar to reported data (Fritsch et al., 2010)]. Elucidation of three different proteins in a single complex is one of the promised capabilities of SiMPull (Jain et al., 2011), and the data we present here constitutes one of the first demonstrations of such a capability.

In order to corroborate our SimPull observations on the existence of ORC-ORCA-H3K9 KMTs and G9a-ORCA-Suv39H1 in a single complex, we utilized sequential IPs. We carried out triple transfections of T7-ORCA, HA-ORC1 and Flag-G9a in U2OS cells, followed by immunoprecipitation of HA-ORC1. Following HA peptide elution, the eluate was used for T7-Ab immunoprecipitation. T7-ORCA was immuprecipitated and a robust co-IP of Flag-G9a was detected (Figure 6C). This further confirmed the existence of ORC-ORCA-H3K9 KMTs in a single complex. Similarly, we performed triple transfections of T7-ORCA, HA-G9a and Flag-Suv39H1 in U2OS cells followed by immunoprecipitation of HA-G9a. Following HA peptide elution, the eluate was used for T7 Ab immunoprecipitation. T7-ORCA was immuprecipitated and a robust co-IP of Flag-Suv39H1 was detected (Figure 6D), further confirming the existence of multiple H3K9 KMTs in a single complex with ORCA. The exogenous expression of Suv39H1 did not affect the association of G9a and ORCA; similarly, the exogenous expression of G9a did not compromise the association of Suv39H1 and ORCA (Figure 2Aa-Ab).

II.2.4. Loss of ORCA causes global changes in H3K9-containing heterochromatin structure

Previous studies indicated that ORCA along with ORC associates with heterochromatic regions (Prasanth et al., 2010; Shen et al., 2010) and also specifically binds to repressive histone and DNA marks (Vermeulen et al., 2010). Since ORCA interacts with both
H3K9me2 and H3K9me3-catalyzing enzymes, we examined the direct binding of ORCA to these marks. We performed peptide pull-downs with N-terminal peptides of histone H3 with the K9 differentially modified with acetylation or mono-, di- or tri-methylation (Figure 8A). Iodoalkyl agarose conjugated peptides were incubated with purified His-tagged-ORCA. We found that ORCA displayed increased interaction with mono-, di- and tri-methylated H3K9, compared to unmodified or K9-acetylated H3 peptides (Figure 8A).

To get a more quantitative estimation of the affinity of ORCA for differentially methylated H3K9, we have employed SiMPull for the first time as a potential substitute for Isothermal Calorimetry. Biotinylated histone H3 N-terminal tails were immobilized on passivated quartz slides followed by passing lysates expressing full-length YFP-ORCA or the fragment 1-127aa which contains only the Leucine-rich repeats and lacks WD domain necessary for binding to methylated histones (Fig. 7Aa). The level of YFP-ORCA expression was quantitated by a direct anti-GFP pull-down with both the lysates (Figure 8Ba) and analyses of the average number of molecules pulled down (Figure 8Bb-Bc). The lysates were then diluted so that the expression of the YFP-tagged proteins is nearly equal and passed through the flow chambers containing the immobilized peptides. Analysis of the average number of molecules pulled down by the peptides revealed that ORCA has the highest affinity for H3K9me3 followed by for me2 and me1 (Fig. 7Ab-c). YFP-ORCA 1-127aa showed a low basal binding to all the peptides corroborating the necessity of WD domain of ORCA for specifically recognizing methylated histones.

To further determine whether ORCA is required for the establishment of these histone marks on chromatin, depletion of ORCA [siRNA-mediated knockdowns (KD)] was carried out both in cancerous cells (U2OS) and diploid fibroblasts (WI-38) and the total levels of H3K9me2 and H3K9me3 were analyzed by immunoblotting (Fig. 4Ba-b). In both the cell lines, upon depletion of ORCA, the levels of H3K9me2 decreased while H3K9me3 remained unchanged at the resolution of western blotting (Fig. 7Ba-b).

To determine the involvement of ORCA in the genome-wide status of H3K9 methylation we conducted H3K9me3 ChIP-sequencing upon ORCA depletion. We observed a significant decrease in H3K9me3 ChIP-seq signal upon ORCA knockdown (Fig. 9A and B). Around 18% of the detected peaks showed highly significant (more than 5 fold
decrease) changes in H3K9 trimethylation in cells lacking ORCA (Fig. 9B, Supplementary File 1 and total H3K9 in Figure 10Aa). Interestingly, several regions did not show significant change in H3K9me3 association upon ORCA KD (Figure 10Ab). Since ORCA associates with heterochromatin, we specifically analyzed the number of reads of satellite repeats in H3K9me3 ChIP and found that there was a significant reduction of this mark at these regions of the genome upon ORCA KD (Fig. 9Ca). Further, H3K9me3 showed less association both with the telomeric (TAR1) and centromeric (SST1) repetitive DNA in cells lacking ORCA (Fig. 9Cb-c).

Our attempts on H3K9me2 ChIP-seq did not succeed because of the technical challenge associated with sequencing the broad H3K9me2 peaks. Similar problems with H3K9me2 ChIP-seq have been previously reported by other studies (Yuan et al., 2009). As an alternate, regions that showed significant reduction of H3K9me3 in the ChIP-seq experiment (as evident by the wiggle plots; Fig. 9Da-d and Figure 10Ba-b) were chosen for H3K9me2 ChIP-qPCR validation. These regions also consistently showed a significant reduction of H3K9me2, corroborating the decrease seen in western blotting (Figure 10Ca-b). *C-FOS*, a gene that does not associate with these repressive histone marks, was used as a negative control (Figure 10Bc and Ca-b). To determine whether the H3K9-targets are directly regulated by ORCA, we conducted ChIP using HA antibody in HA-ORCA expressing stable cell line. This allowed us to address if ORCA is associated with H3K9-occupied genomic sites. We observed a strong enrichment of ORCA at the H3K9-enriched loci (Fig. 9Ea), while ORCA binding to *C-FOS*, a region devoid of H3K9, was comparable to that of IgG (Fig. 9Eb).

To understand the mechanism of reduction of H3K9 methyl marks upon ORCA depletion, we first determined whether the protein stability or chromatin association of G9a and Suv39H1 were altered upon ORCA loss. Our data revealed that the total cellular levels of G9a and Suv39H1 were not reduced upon ORCA KD (Figure 10D). We next addressed if the loading of these KMTs onto chromatin is impaired upon ORCA KD. To investigate this, we performed G9a and Suv39H1 ChIP upon ORCA knockdown and analyzed the association of these enzymes to the loci that show H3K9me2 and H3K9me3 reductions. Suv39H1 showed a decrease at these loci (Fig. 9Fa-b), indicating that the loading of Suv39H1 to these regions is reduced upon ORCA depletion. G9a association
with these regions showed either no alteration or an increase at some regions (Figure 10Ea-b), indicating that the reduction in H3K9me2 that was observed was possibly due to impaired catalytic activity of G9a.

II.2.5. ORCA stabilizes the multimeric H3K9 KMT complex

Next we addressed if ORCA facilitates the assembly of the multimeric H3K9 KMT megacomplex. To address this, the association between the components of the KMT megacomplex, namely G9a and Suv39H1, was evaluated in cells that were treated with control or ORCA siRNAs. Flag-G9a and HA-Suv39H1 were expressed and HA IP was carried out in control and ORCA-depleted (ORCA knockdown KD) cells. ORCA KD showed close to 50% reduction of Suv39H1 that co-immunoprecipitated with G9a (Fig. 11Aa-b). This observation suggested that the stability of the KMT complex requires ORCA. We used SiMPull to obtain an accurate quantitative estimate of the reduction (Fig. 11Ba-c). YFP-Suv39H1 and mCherry G9a were expressed in cells depleted of ORCA. GFP pull down was carried out and the number of mCherry-G9a molecules associated with Suv39H1 was calculated (Fig. 11Bb-c). ORCA knockdown led to ~50% reduction in the complexes containing YFP-Suv39H1 and mCherry-G9a (Note, 24±3% mCherry-G9a pulled down by YFP-Suv39H1 in control versus 15±1% in ORCA knockdown cells; Fig. 11Bc). These results support the argument that ORCA acts as a scaffold protein that is necessary for stabilizing a subset of the complexes containing multiple H3K9 KMTs.

To further confirm the role of ORCA as a scaffold protein, we addressed if over-expressing ORCA leads to any increase in G9a and Suv39H1-containing complexes. We performed triple transfections of YFP-Suv39H1, mCherry-G9a and T7-ORCA. YFP-Suv39H1 SimPull was carried out and the number of mCherry-G9a molecules pulled down was analyzed as a percentage of YFP-Suv39H1 pull-down (Fig. 11Ca-c). The presence of full-length T7-ORCA showed 25±1% association between Suv39H1 and G9A. The mutant T7-ORCA (1-270) that does not interact with either G9a or Suv39H1, when expressed along with G9a and Suv39H1 showed 14±3% of mCherry-G9a in complex with YFP-Suv39H1. By contrast, the other T7-ORCA mutant (128-647) that
interacts with both G9a and Suv39H1 stabilized mCherry-G9a and YFP-Suv39H1 complexes in the cell (29±6%; Fig. 11Cb-c).

These results collectively indicated that ORCA, by acting as a scaffold protein stabilizes the association of multiple KMTs in a single complex. In the absence of ORCA, the integrity of this complex is compromised, leading to the reduction in the KMT-associated enzymatic activity and a subsequent reduction of H3K9me2 and H3K9me3 patterns on chromatin.

II.2.6. Changes in chromatin organization upon ORCA loss affect heterochromatin replication

In general, chromatin at the nuclear periphery is significantly enriched with H3K9me2 whereas H3K9me3 is preferentially enriched around nucleolus (Yokochi et al., 2009). Typically, both of these regions replicate late during S-phase indicating that in general repressive histone marks-containing differentially condensed chromatin influences replication timing and chromatin positioning (Julienne et al., 2013). Therefore, we investigated whether the changes in H3K9me2 and H3K9me3 deposition in specific chromatin regions, upon ORCA depletion, also influences their replication timing. We depleted ORCA in U2OS cells and then synchronized the cells so as to analyze the spatio-temporal regulation of replication during S-phase (Fig. 12A). Samples were collected at 4, 8, 12 h post release from aphidicolin arrest with BrdU pulse-labeling prior to sample collection. This was followed by immunofluorescence to score for cells in early, mid and late S phase of cell cycle (Figure 13A). At 8h and 12h time points post-aphidicolin release, ORCA depletion caused dramatic reduction in cells showing late replication patterns (Fig. 12B). BrdU-PI flow cytometry results showed a significant reduction in BrdU incorporation in ORCA-depleted cells without significant changes in the early S-phase (Fig. 12C). To determine if the changes in late replication pattern are a reflection of changes in the heterochromatin organization, we examined the replication timing of regions that showed a reduction in H3K9me2 and H3K9me3 upon ORCA KD. Initial analysis of the available repli-seq dataset from various human cell lines in UCSC Genome Browser and ENCODE consortium revealed that the replication timing of large domains remains the same across cell lines. We therefore compared the H3K9me3 ChIP-
seq data set to the HeLa repli-seq dataset (Fig. 12D). HeLa-S3 G1b and HeLa-S3 S1 are deep sequencing data sets for late G1 and early S replicating regions in HeLa-S3 cells (Hansen et al., 2010). The chromosomal regions that are replicating at these two stages are shown in black (early) and late (gray) along the length of the chromosome (Fig. 12D and Figure 13B).

Using the dataset mentioned above, we examined the replication timing of the regions that showed reduction in H3K9me3 by ChIP-seq. On chromosome 19, the total H3K9me3 peaks in the control sample and the regions, which show greater than 5-fold decrease in H3K9me3 upon ORCA depletion are represented as black bars above the HeLa-S3 G1b and HeLa-S3 S1 tracks. Upon ORCA depletion, most of the affected H3K9me3 peaks resided in late replicating domains (Fig. 12D and Figure 13B). This coupled with the loss of late replication patterns by BrdU IF in ORCA-depleted cells made us hypothesize that ORCA could also regulate the replication of late replicating regions.

Loss of ORCA could be causing either changes in replication timing of late replicating regions or the complete loss of replication of these regions. To investigate these possibilities, we conducted BrdU ChIP in control and ORCA-depleted U2OS cells (Figure 13Ca-c). We depleted ORCA and arrested the cells using aphidicolin. The cells were then released into S phase, pulse-labeled with BrdU and analyzed by BrdU ChIP at different time points post-release (0, 4, 8 and 12h). The replication timing of various loci that showed significant reduction in H3K9me2 and me3 (Fig. 9D and Figure 10Ca) were assessed by quantitative PCR. We observed changes in replication timing of these loci (Figure 13Cb-c shows two representative loci CELSR3 and FAM20A) upon loss of ORCA. For example, in control cells a significant population of CELSR3 locus replicates in late S (12 hours post release) as evident by the BrdU ChIP signal at 12h. Upon ORCA knockdown, there is a significant increase in the population of the locus replicating during early S (4h timepoint) and a concomitant reduction in BrdU ChIP signal at mid and late S (8 and 12h timepoints) (Figure 13Cb). The replication timing of C-FOS, a region that serves as a control showing no changes in H3K9me2 and me3 upon loss of ORCA, remains unaffected (Figure 13Ca), suggesting that the replication timing changes
observed in ORCA-depleted cells may not be because of the direct role of ORCA/ORC in establishing the pre-replicative complex.

II.2.7. ORCA’s role in heterochromatin organization is independent of its role in preRC assembly

We have previously demonstrated that ORCA plays a key role in replication initiation (Shen et al., 2012). We addressed whether the observed defects in heterochromatin organization and replication patterning in cells lacking ORCA are due to defects in preRC assembly or reflect the more direct role of ORCA in heterochromatin organization. While it is well known that ORC (along with ORCA) associates with heterochromatic regions in post-replicated cells in metazoans (Prasanth et al., 2004; Shen et al., 2010), its direct role in heterochromatin organization versus heterochromatin replication licensing has remained to be understood.

In order to understand ORCA’s role in chromatin organization and if it is independent of its role in preRC function, we wanted to deplete ORCA after the establishment of pre-replication complex (post G1 phase) but before DNA synthesis began. Depletion of a protein within a short, specific time window by RNA interference is challenging because even if the mRNA levels are dramatically reduced, the protein levels could persist for significantly longer times. This necessitates the use of a strategy that utilizes post-translation degradation process for reducing proteins levels efficiently. To achieve this, we utilized a commercially available Proteotuner kit (Clontech). Briefly, an siRNA resistant version of T7-ORCA (T7-ORCA-siRNA NTV: Non-Targetable Version) was tagged with a destruction signal or DD (Destabilization Domain) tag, a destabilization domain of the FKBP protein [Fig. 14A; (Banaszynski et al., 2006)]. This signal is recognized by the proteosomal machinery and results in the rapid degradation of the exogenous ORCA. In the presence of a ligand, Shield1, the DD tag is masked by the binding of Shield1, thereby preventing the degradation of the exogenous ORCA protein.

In order to determine whether DD-T7-ORCA-siRNA NTV is functional and can substitute for endogenous ORCA, we examined whether it could efficiently rescue ORC levels on chromatin upon depletion of endogenous ORCA (Shen et al., 2010). We transfected DD-T7-ORCA-siRNA NTV into U2OS along with siRNA to knockdown
endogenous ORCA. Following two rounds of siRNA treatment (48 h) in the presence of DD-T7-ORCA-siRNA NTV (Fig. 14B), we examined the loading of ORC2 on chromatin and compared it to ORC2 loading in control and ORCA-depleted cells. We observed that while ORC2 loading was affected upon ORCA depletion (levels of ORC2 decrease in the chromatin fraction with concomitant increase in the supernatant fraction), the expression of DD-T7-ORCA-siRNA NTV construct rescued this phenotype by restoring the levels of ORC2 on chromatin to an extent comparable to that of the control (Fig. 14C). In addition, we also carried out immunoprecipitation of DD-T7-ORCA and found that it efficiently interacts with endogenous ORC2 (Figure 15A), further confirming that DD-T7-ORCA is functional.

To determine if the role of ORCA in heterochromatin organization is independent of its role in preRC assembly, we transfected U2OS cells with DD-T7-ORCA-siRNA NTV. We then depleted endogenous ORCA by using ORCA siRNA while the levels of exogenous DD-T7-ORCA-siRNA NTV were maintained by growing the cells in the presence of Shield1. We synchronized the cells at early S by using aphidicolin and then degraded exogenous DD-T7-ORCA at early S by removing shield from the medium. Removal of Shield1 resulted in the loss of exogenous ORCA [in addition to endogenous ORCA that was removed by siRNA treatment (Fig. 14D)]. The cells were then allowed to progress through S phase and chromatin organization and replication were examined at different time points during S phase. Specific depletion of ORCA only in post-G1 cells also resulted in reduction in the H3K9me2 levels (Fig. 14D). This demonstrates that the heterochromatin is disorganized in the absence of ORCA in post-G1 cells.

In these cells we examined the replication patterning by BrdU immunofluorescence. We observed a decrease in cells showing mid- and late- patterns of DNA replication and a concomitant increase in cells showing early patterns (Fig. 14Eb), similar to our previous observations (Fig. 12B). In cells lacking ORCA, a large number of cells showing mid-replication patterns showed BrdU staining preferentially at perinucleolar regions (Fig. 14Ea, Figure 15B). Furthermore, we observed that H3K9me3 and HP1α were mislocalized and formed perinucleolar rings in cells lacking ORCA (Fig. 14Fa-b, Figure 15C). Such localization was reminiscent of HP1α localization in Orc1 and Orc5-depleted
cells (Prasanth et al., 2010). Moreover, both control and ORCA-depleted cells progressed through S phase at comparable rates (Figure 15D) indicating that the observed defects in HP1α localization and BrdU incorporation upon loss of ORCA are due to defects in heterochromatin and not due to indirect effects of defects in S phase progression. Based on the results, we propose that the observed defects in heterochromatin organization in ORCA-depleted post-G1 cells are independent of its known functions in preRC assembly.

II.3. Discussion

ORCA, a key player in the initiation of DNA replication, associates with multiple components of the pre-replicative complex (Shen et al., 2012). The ORCA-ORC complex associates with heterochromatin, including telomeric and centromeric regions even after replication has been accomplished suggesting that ORCA-ORC complex may play key roles in heterochromatin organization in addition to its role in pre-RC. The WD-repeat-containing domain (also found in ORCA) mediates interaction of proteins with nucleosomes/histones (Suganuma et al., 2008). For example, WDR5, a component of the MLL/SET1 KMT complex binds to H3K4me2 using WD repeats (Ruthenburg et al., 2006). Similarly, HIRA, a WD-repeat-containing protein, binds to core histones and controls transcription (Lorain et al., 1998). Here, we demonstrate that ORCA associates with multiple H3K9 KMTs, binds to methylated H3K9 and regulates both the organization and replication of repressed chromatin marked with H3K9me2 and H3K9me3. Recently, a H3K9 KMTs multimeric complex has been identified, that has been shown to be recruited to major satellite repeats and a subset of promoters of G9a-repressed genes and a functional cooperation of these enzymes is crucial for the regulation of G9a target genes (Fritsch et al., 2010). We demonstrate that ORCA-ORC associates with the H3K9 KMT-containing complex and in the absence of ORCA, this complex disintegrates. The loss of the enzymatic activity of this complex causes changes in the H3K9me2 and H3K9me3 profile in human cells. Based on this, we propose that ORCA is a scaffold protein that stabilizes the H3K9 KMT complex.

Recent work suggests that in mouse cells ORCA associates with pericentric heterochromatin via its association to H3K9me3 and maintains silencing at the major
Based on our results, we speculate that the changes in transcription of satellite repeats upon ORCA-depletion are likely caused by the changes in the heterochromatin structure.

Immunoprecipitation experiments demonstrate that the WD-repeat of ORCA and the ankyrin repeat of G9a are crucial for the interaction between these two proteins. Ankyrin-repeats of G9a also contain methyl-lysine binding modules and can therefore generate as well as read the same epigenetic mark (Collins et al., 2008). We have observed that the SET domain of G9a and its catalytic activity is essential for the binding of ORCA to heterochromatin, suggesting that the chromatin modifications initiated by the KMTs provide docking sites for ORCA. These in turn may facilitate the recruitment of accessory factors that stabilize the interaction and help the propagation of heterochromatin. We propose that ORCA recognizes repressive histone marks, binds to KMTs that in turn facilitate the propagation of the histone mark. The newly established marks then become docking sites for ORCA and the whole process is repeated and this results in the spreading of heterochromatin (Fig. 16A).

Trimethylation of H3K9, monomethylation of H3-K27 and trimethylation of H4-K20 are enriched at pericentric heterochromatin (Peters et al., 2003; Rice et al., 2003; Schotta et al., 2004). It is well established that H3K9 trimethylation is a prerequisite for the subsequent H4K20 trimethylation at the pericentric heterochromatin and this in turn sets the chromatin for binding of other key heterochromatin proteins including HP1 (Fischle et al., 2005; Lachner et al., 2001; Stewart et al., 2005). It is interesting to note that ORCA also interacts with Suv420H1 and H2, enzymes that catalyze H4K20 di and trimethylation respectively (data not shown). It has been previously reported that Suv420H2 is a structural component of the heterochromatin and is required for chromatin compaction as well as cohesin recruitment (Hahn et al., 2013). Recently, Reinberg and coworkers have proposed that the H4K20 me1/2/3 is also crucial for the regulation and timing of replication origin firing and that ORCA and Orc1 directly recognize these chromatin sites (Beck et al., 2012). We are currently addressing the functional relevance of ORCA and Suv420H1/2 interaction in heterochromatin organization and replication progression.
Work from Jenuwein and co-workers have pointed to the idea that Prdm3 and Prdm16, H3K9 mono-methyltransferases, are also required for mammalian heterochromatin formation (Pinheiro et al., 2012). Similarly, mono-methylation of H3 at K9 catalyzed by SETDB1 has been shown to be a favored substrate for Suv39h for K9 trimethylation, which can then establish heterochromatin (Loyola et al., 2009). The mechanism by which the Suv39h or H3K9me3 is targeted to pericentromere has been a long-standing question. It is generally assumed that HP1 is a key regulator of heterochromatin organization that is required for establishment and maintenance of this compacted form of chromatin. Spreading of the heterochromatin is thought to involve a mechanism where HP1 dimerizes, interacts with Suv39h1/2 and also recruits de novo DNA methyltransferase activity (Almouzni and Probst, 2011). The fact that HP1 associates with heterochromatin in a transient manner has suggested that other perhaps constitutively bound factors contribute to the organization of heterochromatin (Cheutin et al., 2003). In addition recent work has demonstrated that pericentric heterochromatin can be generated independent of Suv39h-HP1 binding (Muramatsu et al., 2013). A transcription factor-based mechanism has also recently been suggested as an intrinsic mechanism for the formation of heterochromatin in mouse cells (Bulut-Karslioglu et al., 2012). Based on our results, we propose that in human cells ORCA facilitates the recruitment, accumulation and also the propagation of the heterochromatin by a self-sustaining loop mechanism, whereby ORCA binds to specific chromatin marks, associates with Suv39H1, that in turn propagates more H3K9me3 marks, generating more docking sites for ORCA (Fig. 15).

We previously demonstrated that ORCA can facilitate the binding of ORC to chromatin and in the absence of ORCA, the binding of ORC to chromatin is compromised (Shen et al., 2010). However, it remains to be determined if the loss of chromatin-bound ORC in ORCA-depleted cells occurs at specific origins or at heterochromatic sites or both. Our data unequivocally demonstrates that ORCA binds to H3K9 methylated chromatin and facilitates the recruitment of KMTs as well as other components of ORC and MCMs to these sites. It is possible that the regulation of these epigenetic modifications by ORCA may provide identity to repressed chromatin and this in turn is crucial for proper replication. This idea is supported by our observation that ORCA associates with
repressive KMTs only in a heterochromatic environment. It is known that G9a mediates dimethylation of H3K9 at late-replicating chromatin and this occurs predominantly at the nuclear periphery. H3K9me2 and H3K9me3 are enriched at the late-replicating facultative or constitutive heterochromatin respectively (Wu et al., 2005). The reduction of these marks upon loss of ORCA leads to changes in the replication timing only of these regions as indicated by the significant decrease in late replication patterns upon ORCA knockdown. This is very similar to previous reports that show that loss of Rif1 causes reduction in mid-replication patterns and global changes in replication timing primarily due to Rif1’s role in organizing chromatin loops (Cornacchia et al., 2012; Kumar and Cheok, 2014; Yamazaki et al., 2012). ORCA could be functioning in a similar fashion as an organizer of heterochromatin and therefore in cells lacking ORCA replication timing is altered. Furthermore, using the proteotuner system, we demonstrate that the role of ORCA in chromatin organization is independent of its role in preRC assembly.

Heterochromatin is formed as a result of a maturation process that requires several steps and ORCA acts early in this process. Our results demonstrate that ORCA is a chromatin reader that facilitates the assembly of the writer KMT complex and its associated partners to specific chromatin sites. These together regulate key cellular events, including DNA replication and heterochromatin organization.

II.4. Materials and methods:

II.4.1. Plasmid constructs:

Human G9a (hG9a) full length and mutant clones were obtained using PCR using pCDNA3 Flag G9a plasmid provided very kindly by Dr. Martin Walsh. The mutants were cloned into pCGN, pEYFP, pECFP and pEmCherry vectors (clonetech) and pEYFP-LacI vector. pEGFP G9a full length and ΔSET constructs were also kindly provided by Dr. Walsh. Mouse G9a (mG9a) full length was amplified from pSV2 YFP mG9a (Dr. David Spector’s lab) (Janicki et al., 2004). pEYFP LacI mG9a ΔSET and H1166K constructs were cloned by amplification from respective pSG5 HA mG9a constructs kindly provided by Dr. Michael Stallcup. Flag GLP was kindly provided by Dr. Dan Levy.
Human Suv39H1 full length and mutant clones were obtained using PCR using Flag Suv39H1 plasmid provided very kindly by Dr. Rama Natarajan (Villeneuve et al., 2008). The mutants were cloned into pCGN vector and pEYFP vector (clonetech). pSV2-YFP-mSuv39H1 has been described previously (Janicki et al., 2004).

Myc-EZH2 was kindly provided by Dr. Francois Fuks, Free University Brussels. EZH2 was PCR amplified and cloned into pEYFP-LacI vector.

Flag Suv420H1.1 and H2 constructs were kindly provided by Dr. Craig Mizzen.

pEGFP-LacI vector was a kind gift from Dr. Miroslav Dundr (Kaiser et al., 2008) and used for making pEYFP-LacI vector. T7 ORCA mutants, pEYFP and CFP ORCA, pECFP-LacI and pECFP-LacI-ORCA have been described previously (Shen et al., 2010). YFP-LacI-Orc2 was cloned by amplifying and inserting Orc2 into pEYFP-LacI vector.

YFP-Orc1 construct has been described previously (Hemerly et al., 2009).

T7-ORCA-siRNA NTV was cloned into pPTuner IRES2 (Clonetech) vector.

All the cloned constructs were confirmed by sequencing and used for immunoprecipitation and/or cell biological experiments.

**II.4.2. Cell culture:**

U2OS cells were grown in Dulbecco’s modified Eagle’s medium (DMEM) containing high glucose and supplemented with 10% fetal bovine serum (FBS – Hyclone). WI38 cells were also grown in the same medium but supplemented with non-essential amino acids. Hela suspension cells (Hela-XLP GLP) were grown in DMEM supplemented with 5% fetal calf serum and penicillin-streptomycin. U2OS-2-6-3 CLTon cells were grown in DMEM with 10% Tet system approved FBS (Clonetech).

For arresting cells at G1/S, Aphidicolin (stock 10mg/ml) was added to the cells at a final concentration of 5µg/ml for 12 h. Cells were then washed 3 times with PBS and released into S phase with medium (DMEM + 10%FBS) lacking aphidicolin. Cells were then collected at 4, 8 and 12h post G1/S block for immunofluorescence and flow cytometry analysis.

siRNA transfection for ORCA depletion: Cells were grown to 30% confluency and siRNA against ORCA or control luciferase gene (Shen et al., 2010) were delivered into
the cells at a final concentration of 100nM using Lipofectamine RNAimax (Invitrogen). The siRNAs were delivered twice at the gap of 24 hours and the cells were collected 24 hours after the second round of transfection for subsequent analysis.

Rescuing of ORC loading by using DD-T7-ORCA-siRNA NTV:

U2OS cells were transfected with of DD-T7-ORCA-siRNA NTV construct along with siRNA against ORCA. Five hours later Shield1 (0.5µM) was added to the medium. 24 hours after the first round of knockdown, a second round of ORCA siRNA treatment was carried out in the presence of Shield1. Samples were collected 24h later for chromatin fractionation to examine ORC loading.

Depletion of ORCA using proteotuner:

As described above, DD-T7-ORCA-si NTV (500ng plasmid was transfected) was expressed in U2OS cells grown on coverslips in the presence of Shield1 (0.25µM). This was followed by addition of fresh medium containing aphidicolin (5µg/ml) + Shield1 (0.25µM). 14h post aphidicolin block, cells were washed thrice with PBS containing aphidicolin, with or without Shield1 respectively. This was followed by performing control and ORCA depletions in Optimem containing aphidicolin, with or without Shield1. The control and ORCA-depleted cells were washed with PBS containing or lacking Shield1 respectively. Then the cells were released into S phase using medium with or without Shield1 for control and ORCA-depleted cells followed by late S sample collection 12h later for western blotting and immunofluorescence analysis.

II.4.3. Insect cell culture and baculovirus expression:

His ORCA, G9a and Suv39H1 viruses was generated by using pFastBac HT-B-His tagged ORCA, G9a and Suv39H1 respectively (Shen et al., 2012) (by following Bac-to-Bac baculovirus expression system – Invitrogen). Virus production was carried out in Sf9 cells with viruses collected 72 hours post infection (multiplicity of infection 5 to 10). Protein expression was carried out in Hi5 cells by collecting cells 65 hours post infection. Nuclei were collected by using Hypotonic lysis buffer (20mM potassium phosphate buffer pH 7.5, 5mM KCl, 1.5mM MgCl2 and 5mM b-mercaptoethanol) making nuclear extracts in PK50 buffer (20mM potassium phosphate buffer pH 7.5, 50mM KCl, 0.02%
NP-40, 10% Glycerol, 5mM b-mercaptoethanol with protease and phosphatase inhibitors) (Siddiqui and Stillman, 2007). 45% Ammonium Sulphate precipitation was carried out followed by reconstitution of His-ORCA, G9a and Suv39H1 in PK50 buffer.

**II.4.4. Immunofluorescence:**

Cells were fixed with 2% Formaldehyde in phosphate buffered saline (PBS -pH 7.4) for 15 minutes in room temperature followed by permeabilization with 0.5% Triton X-100 in PBS for 7 minutes on ice or pre-extracted before fixing with 0.5% Triton X-100 in Cytoskeletal buffer (CSK: 100mM NaCl, 300mM Sucrose, 3mM MgCl₂, 10mM PIPES at pH 6.8) for 5 minutes on ice. Blocking was then done for 30 minutes with 1% Normal goat serum (NGS) in PBS. Primary antibody incubation was then carried out for 1 hour in a humidified chamber followed by secondary antibody incubation for 25 minutes. The cells were then stained with DAPI and mounted using vectashield (Vector Laboratories Inc.). The following antibodies were used for immunofluorescence: BrdU (1:500; Sigma mAb BU-33), Lamin (1:750), H3K9me2 (1:100; Millipore 07-212), H3K9me3 (1:200, Millipore 07-523), HP1α (1:100, Millipore 3584).

**BrdU Immunofluorescence:** After primary and secondary antibody incubation for lamin immunofluorescence, cells were fixed with 2% Formaldehyde solution in PBS. This was followed by acid denaturation of DNA using 4N HCl for 25 minutes at room temperature. Three washes with PBS and two washes with PBS-NGS followed. This was followed by incubation of the cells with Anti-BrdU antibody followed by secondary antibody incubation and mounting.

Zeiss Axioimager z1 fluorescence microscope (Carl Zeiss Inc.) equipped with chroma filters (Chroma technology) was used for observing the cells and statistics. Axiovision software (Zeiss) was used for digital imaging using Hamamatsu ORCA cooled CCD camera. Cells were also examined on the Delta vision optical sectioning deconvolution instrument (Applied precision) on an Olympus microscope.

**II.4.5. Immunoprecipitations and immunoblots:**

For co-immunoprecipitation with transiently transfected HKMTs and ORCA, co-transfections were carried out in U2OS cells. Cells were lysed, 24 hours post-
transfection, in IP buffer (50mM Tris pH 7.4, 150mM NaCl, 10% glycerol, 0.1% NP-40, 1mM DTT supplemented with the protease and phosphatase inhibitors). After pre-clearing with Gammabind Sepharose beads for 1 hour, the lysates were incubated with appropriate antibody-conjugated agarose overnight. The beads were washed in the same IP buffer and finally denatured by the addition of Laemmli buffer. The complex was analyzed by Western blotting.

For immunoprecipitations and immunoblotting the following antibodies were used anti-GFP (1:500; Covance), anti-Flag M2 (1:500, Sigma), anti-HA 12CA5 (1:100) and anti-T7 (1:5000; Novagen), anti-ORCA pAb 2854-1 AP (1:500), anti-G9a (1:500, Sigma), anti-Suv39h1 (1:200, Cell Signaling), anti-ORC2 pAb 205-6 (1:1000), anti-Geminin (1:1000, Santa Cruz), anti-MCM3 TB3 (1:750), anti-α-tubulin (1:10,000, Sigma-Aldrich), anti-H3K9me2 (1:200, Millipore 07-212), anti-H3K9me3 (1:500, Millipore 07-523), anti-SF2 AK96 (1:750), anti-PCNA mAb PC10 (1:750) antibodies.

For IP in the presence of EtBr, lysates were made with IP buffer described above followed by addition of EtBr (stock 10mg/ml, working 50µg/ml). After pulldown, beads were washed with IP buffer supplemented with 80 µg/ml of EtBr. For Benzonase treatment cells (grown in 6cm plates) were lysed for 10 min in IP buffer (50mM HEPES pH 7.9, 10% Glycerol, 200mM NaCl, 0.1% Triton X-100,1mM MgCl₂) supplemented with protease and phosphatase inhibitors. 1000 U of Benzonase (Sigma) was then added followed by nutation at room temperature for 20 min. EDTA (final concentration 5mM) was added to stop the reaction followed by centrifugation at 12500 rpm, 5 min at 4°C. The supernatant was used for subsequent antibody incubation.

II.4.6. Nuclear extracts for semi-endogenous IP:

The nuclear extraction protocol has been described previously (Fritsch et al., 2010; Robin et al., 2007). We carried out HA immunoprecipitation in Hela cells stably expressing Flag-HA-GLP and Flag-HA-G9a stable by retroviral transduction. First, nuclear extract was made using an equivalent of 20g of dry cell pellet, which approximately corresponds to 3 billion cells. Cells were resuspended in Hypotonic buffer (10mM Tris pH 7.6, 1.5mM MgCl₂, 10mM KCl) with the volume of hypotonic buffer being equal to the packed volume of cells. The suspension was then dounced 20 times using tight pestle
followed by adding one third the packed volume of sucrose buffer (20mM Tris pH 7.6, 15mM KCl, 60mM NaCl, 0.34M Sucrose, 0.15mM Spermine, 0.5mM Spermidine). Then centrifugation was carried out (9000rpm, 7 min). The supernatant was discarded and the nuclei were resuspended in low salt buffer (20mM Tris pH 7.6, 25% glycerol, 1.5mM MgCl$_2$, 0.2mM EDTA, 20mM NaCl). This was followed by adding high salt buffer (20mM Tris pH 7.6, 25% glycerol, 1.5mM MgCl$_2$, 0.2mM EDTA, 900mM NaCl) dropwise while vortexing to make the final salt concentration 300mM. After incubation on ice for 30 min, one third the volume of sucrose buffer was added followed by centrifugation at 1000rpm, 10 min, 4°C. The supernatant is the nuclear soluble fraction. The pellet (chromatin bound fraction) was resuspended in sucrose buffer and CaCl$_2$ (final concentration 1mM) was added. The sample was then preheated for one min at 37°C and MNase was added to a final concentration of 0.0025U/µl. The sample was then incubated for exactly 12 min at 37°C followed by placing the tubes on ice. 0.5M EDTA was added at a final concentration of approximately 3.6µM. The samples were then sonicated (Bioruptor, high amplitude 5cycles: each cycle 1min ON, 1min OFF). Protease and phosphatase inhibitors were added and the samples (nuclear soluble and chromatin bound fractions) were ultracentrifuged at 40000rpm for 30 min. The supernatants were transferred to a fresh tube. Tagged-H3K9 HMTs complexes were then purified by immunoprecipitation using anti-FLAG antibody immobilized on agarose beads (cat# A2220, Sigma). After elution with the FLAG peptide (Ansynth, The Netherlands), the bound complexes containing nucleosomes were further affinity-purified on anti-HA antibody-conjugated agarose (cat# A2095, Sigma) and eluted with the HA peptide (Ansynth, The Netherlands). The samples were then analyzed by immunoblotting.

**II.4.7. Nuclear extracts for Single-Molecule Pull-down:**

Cells (Hela for gel filtration and U2OS for single-molecule pull-down) were lysed in hypotonic buffer (10mM HEPES-NaOH pH7.9, 10mM KCl, 2mM MgCl$_2$, 0.34M Sucrose, 10% glycerol, 0.1% Triton X-100) supplemented with 1mM DTT, protease and phosphatase inhibitors. The lysate was incubated at 4°C for 5 min after which nuclei were collected by centrifuging at 1500g for 5 min. The pellet was then resuspended in Nuclear extraction buffer (10mM HEPES-NaOH pH7.9, 2mM MgCl$_2$, 1mM EGTA, 25% Glycerol, 350mM NaCl, 0.1% Triton X-100 and 1mM DTT) supplemented with protease
and phosphatase inhibitors. The supernatant was collected by centrifugation at 12000rpm for 5min. The lysate was then used for co-localization studies by SiMPull.

**II.4.8. Direct interaction studies:**

Baculovirally purified His-ORCA and His-G9s/His-Suv39H1 were diluted using PK 50 buffer and incubated together for 2h at 4°C in the presence of ORCA antibody or pre-bleed. ORCA containing complexes were then pulled down followed by washes with PK 150 buffer (20mM potassium phosphate buffer pH 7.5, 150mM KCl, 0.5% NP-40, 10% Glycerol, 5mM b-mercaptoethanol with protease and phosphatase inhibitors). The samples were finally denatured by the addition of Laemmli buffer. The complexes were analyzed by Western blotting.

**II.4.9. Flow cytometry – BrdU-PI staining:**

U2OS cells were grown in 6cm plates to approximately 50% confluency followed by incorporation of BrdU (stock 10mM; working 50µM) for 1h at 37°C. Cells were then harvested at 3500 rpm, 15 min followed by washing with 1% BSA in PBS (pH 7.4). The cells were then resuspended in 0.9% NaCl (final cell density: 2 X 10^6 cells/ml). The cells were then fixed by adding chilled 100% Ethanol to a final concentration of 50% (left overnight at -20°C). After spinning down the fixed cells, DNA was denatured by resuspending in 2N HCl + 0.5% Triton X – 100 and incubating for 30 min at room temperature. The cells were then pelleted and resuspended in 0.1M Sodium tetraborate pH 8.5. This was followed by centrifugation at 3500 rpm, 15 min at 4°C followed by resuspending the cells in PBS + 1% BSA + 0.5% tween 20. Anti-BrdU FITC antibody (1ug Ab/10^6 cells) was added and the cells were incubated at RT for 1h. 1ml of PBS + 1% BSA + 0.5% tween 20 was added after that followed by spinning down the cells and proceeding with RNase A treatment and PI staining as described in the previous section.

**II.4.10. Peptide pulldown:**

Human Histone H3 (amino acids 1-15) peptides were synthesized (Biomer technology) with a Cysteine at the N terminus. The K9 (Lysine at position 9) of the peptides was either unmodified, acetylated, mono, di or tri methylated. The peptides were dissolved in
water, quantitated using reverse phase chromatography, lyophilized and stored at -20°C as 1mg aliquots.

For coupling the peptides to SulphoLink Coupling Resin (Thermo Scientific), the peptides were reduced first. For this, 1mg of each peptide was dissolved in 3ml of Coupling Buffer (50mM Tris pH 8.7, 5mM EDTA, final pH adjusted to 8.5) supplemented with TCEP HCl (Thermo Scientific) and allowed to incubate at room temperature for 1h. 2ml of the beads were washed with 5ml coupling buffer (three 10min washes) and resuspended in 1ml of coupling buffer. 3ml of the reduced peptides was added to the slurry followed by mixing immediately to distribute the peptide throughout the slurry. The mixture was incubated overnight at room temperature with gentle mixing. The beads were then spun down (2000 rpm, 2min), washed with 6ml of coupling buffer (three 5min washes), resuspended in 5ml coupling buffer + 1ml L-Cysteine HCl. The mixture was incubated overnight at room temperature. The beads were then spun down and washed with 5ml of 1M NaCl (three 5min washes). This was followed by washing two 5min washes with 5ml of storage solution (0.05% NaN3 in water) and final suspension in 5ml of storage solution and storage at 4°C.

50µl packed volume of beads (300ul of bead slurry) was washed with PK 150 buffer (20mM Potassium Phosphate buffer pH 7.5, 150mM KCl, 0.02% NP-40, 10% glycerol, 5mM β- Mercaptoethanol) and incubated with baculovirally purified His-ORCA for 2h at room temperature. This was followed with 5 washes with PK 150 buffer. The beads were then resuspended in Laemmli buffer and analyzed by western blotting.

II.4.11. Single Molecule Pulldown:

SiMPull experiments were carried out in flow chambers prepared on quartz microscope slides which were passivated with methoxy-polyethylene glycol (mPEG) doped with 1% biotin-PEG (Lysan Bio, Inc) (Jain et al., 2011). Appropriate biotinylated-antibody was immobilized on PEG passivated surfaces at approximately 20nM concentration for 20 minutes after coating the flow chambers with 0.2mg/ml NeutrAvidin for 5 minutes. Antibodies were immobilized on NeutrAvidin (Thermo) coated flow chambers either by incubating with biotinylated T7 antibody (Novagen) for 10 min. RIPA buffer lysed samples were then incubated in the chamber for 20 min and washed twice with the buffer
(10mM Tris-HCl pH 8.0, 50mM NaCl 0.1mg/ml BSA). Single-molecule data were acquired by a prism type total internal reflection fluorescence (TIRF) microscope and analyzed using scripts written in Matlab. For ORCA-Orc1; ORCA-G9a; ORCA-Suv39H1 SiMPull analysis, lysates were made from the cells transiently transfected with T7-ORCA with YFP-Orc1, YFP-G9a or YFP-Suv39H1 respectively. For multimeric complex assembly analysis using SiMPull, cells were transfected with T7-ORCA, YFP-Orc1 and mCherry-G9a or T7-ORCA, YFP-Suv39H1 and mCherry-G9a.

For peptide pulldown experiments, biotinylated peptides were immobilized instead of antibodies. Cells lysed in RIPA buffer or Nuclear extracts (depending on the experiment) were then incubated in the flow chamber for 20 minutes followed by wash with T300 buffer (20mM Tris-HCl, pH 8.0, 300mM NaCl, 0.1mg/ml bovine serum albumin [BSA]). Single molecules were visualized by prism-type total internal reflection fluorescence (TIRF) microscope and analyzed using MATLAB scripts (https://github.com/vasuagg/SiMPull_Analysis). Cell lysate was appropriately diluted in T300 buffer to obtain optimal single molecule density on the surface.

II.4.12. SiMPull Data Analysis

Single molecule data was acquired as the average number of YFP or mCherry fluorescent molecules per imaging area (5000µm²) as shown in the histograms. The error bars represent standard deviation of the mean values from 20 imaging areas. Number of fluorescence photobleaching steps was determined for each YFP-tagged molecule and accumulated to obtain the stoichiometry of the complex. Colocalization percentage between YFP and mCherry was calculated as the number of coaligned molecules of one fluorescent molecule with respect to the fluorescent molecules found in lower density on the surface. This was needed since the number of YFP and mCherry tagged proteins were not pulled down to the same extent due to their independent interaction with ORCA. Colocalization criterion was set at 2 pixels, which correspond to a diffraction limited spot (~300nm) for our TIRF setup. Error bars represent standard deviation of the mean values obtained from 3 independent experiments. For pulled down experiments performed using H3K9 peptides, the expression level of YFP-WT ORCA and YFP-mutant-ORCA was compared in the beginning by performing a direct pulldown by anti-GFP. The peptides
pulldown was then performed at appropriate lysate dilution such that protein expression was same for WT and mutant ORCA.

**II.4.13. Chromatin Immunoprecipitation**

H3K9me2 and me3 ChIPs: Formaldehyde (Sigma) was added to culture medium to a final concentration of 1%. Crosslinking was allowed to proceed for 10 min at room temperature and stopped by addition of glycine at a final concentration of 0.125 M. Fixed cells were washed and harvested with PBS. Chromatin was prepared by two subsequent extraction steps (10 min at 4°C) with Buffer 1 (50 mM Hepes/KOH pH 7.5; 140 mM NaCl; 1 mM EDTA; 10% Glycerol; 0.5% NP-40; 0.25% Triton) and Buffer 2 (200 mM NaCl; 1mM EDTA; 0.5mM EGTA; 10 mM Tris pH 8). Nuclei were then pelleted by centrifugation, resuspended in Buffer 3 (50 mM Tris pH 8; 0.1% SDS; 1% NP-40; 0.1% Na-Deoxycholate; 10 mM EDTA; 150 mM NaCl) and subjected to sonication with Bioruptor Power-up (Diagenode) yielding genomic DNA fragments with a bulk size of 150-300 bp. Chromatin was precleared with Protein A/G ultralink beads (53133, Pierce) for 2h at 4°C and immunoprecipitation with the specific antibodies carried out overnight at 4°C. Immune complexes were recovered by adding pre-blocked protein A/G ultralink beads and incubated for 2 h at room temperature. Beads were washed twice with Low salt buffer (0.1% SDS; 1% Triton; 2 mM EDTA; 20 mM Tris pH 8; 150 mM NaCl), twice with High salt buffer (0.1% SDS; 1% Triton; 2 mM EDTA; 20 mM Tris pH 8; 500 mM NaCl), once with LiCl wash buffer (10 mM Tris pH 8.0; 1% Na-deoxycholate; 1% NP-40, 250 mM LiCl; 1 mM EDTA) and twice with TE + 50mM NaCl. Beads were eluted in TE + 1% SDS at 65°C and cross-link was reversed O/N at 65°C. The eluted material was phenol/chloroform-extracted and ethanol-precipitated. DNA was resuspended in water and q-PCR performed using PowerSYBR Green PCR Master mix (Applied Biosystems) and analyzed on a 7300 PCR System (Applied Biosystems). ChIP-qPCR results were represented as percentage (%) of IP/input signal (% input). HA-ORCA ChIPs were carried out using HA-ORCA stable cell lines in U2OS using a similar protocol with the following modifications. All the washing steps after immune complexes pulldown were done once followed by two washes with TE. Beads were eluted with 1% SDS+0.1M NaHCO3 at 65°C and cross-link was reversed O/N at 65°C. The eluted material was
purified using Qiagen gel purification kit and qPCR carried out. G9a and Suv39H1 ChIPs were carried out using double crosslinking protocols. The first crosslinking was carried out using Disuccinimidyl glutarate (DSG) (Santa Cruz; stock 50mM DSG in DMSO) and the second crosslinking using formaldehyde. U2OS cells were grown in 10cm plates to 80% confluency, washed twice with PBS (pH 7.4). Freshly made crosslinking solution (2mM DSG + 1mM MgCl₂ in PBS-pH 8.0) was added for 45 min at room temperature (RT). The cells were then washed twice with PBS (pH 7.4) and 10ml of freshly made crosslinking solution (1% Formaldehyde, 15mM NaCl, 150µM EDTA, 75µM EGTA, 15µM HEPES pH 7.9) was added for 10 min at RT. Then 3ml of freshly made 1M Glycine was added for 5min at RT followed by two cold washes with PBS (pH 7.4). The cells were then pelleted in PBS (supplemented with protease inhibitors) followed by lysis with 300ul of SDS lysis buffer (1% SDS, 10mM EDTA, 50mM Tris pH 8.0). The lysate was then subjected to sonication with Bioruptor Power-up (Diagenode). Chromatin was precleared with Dynabeads protein G (Life technologies) for 2h at 4°C and immunoprecipitation with the specific antibodies carried out overnight at 4°C. Immune complexes were recovered by adding pre-blocked Dynabeads (1mg/ml BSA, 0.4 mg/ml salmon sperm DNA) and incubated for 2 h at 4°C. Beads were washed once with Low salt buffer (0.1% SDS; 1% Triton; 2 mM EDTA; 20 mM Tris pH 8; 150 mM NaCl), once with High salt buffer (0.1% SDS; 1% Triton; 2 mM EDTA; 20 mM Tris pH 8; 500 mM NaCl), once with LiCl wash buffer (10 mM Tris pH 8.0; 1% Na-deoxycholate; 1% NP-40, 250 mM LiCl; 1 mM EDTA) and twice with TE (10mM Tris pH 8.0, 1mM EDTA). Beads were eluted in Elution buffer (1% SDS, 0.1M Sodium Bicarbonate in water) at 65°C twice, 10 min each. The eluates were pooled (250ul), NaCl added (final concentration 0.2M) and cross-link was reversed O/N at 65°C. The eluted material was Rnase A treated (10 µg/ml, 1h at 37°C) followed by Proteinase K treatment (4ul 0.5M EDTA, 8ul 1M Tris pH 6.9, 1ul Proteinase K 20mg/ml) at 42°C for 2h. DNA was purified using QIAquick PCR purification kit (Qiagen) and q-PCR performed SYBR Green PCR Master mix and analyzed on a 7300 PCR System (Applied Biosystems). ChIP-qPCR results were represented as percentage (%) of IP/input signal (% input).
II.4.14. BrdU ChIP after ORCA knockdown and data analysis:

Two rounds of ORCA knockdown were carried out, 24h apart. The cells were then arrested using aphidicolin for 12h followed by release into S phase and samples were collected 0, 4, 8 and 12h post release for BrdU ChIP. Prior to each time point collection, cells were pulsed for 2h with BrdU (10µM).

Cells for each time point were then lysed with 300ul of SDS lysis buffer (1% SDS, 10mM EDTA, 50mM Tris pH 8.0). The lysates were subjected to sonication with Bioruptor Power-up (Diagenode). 100ul sheared chromatin aliquots were then placed on 95°C heat block for 10 minutes. This was followed by snap chilling the samples for 10 minutes. The samples were then diluted, precleared and processed further in a manner identical to the ChIP protocol described in the previous section.

qPCRs were carried out with purified DNA of input, BrdU ChIP and mouse IgG ChIP samples obtained at 0, 4, 8 and 12h post aphidicolin release timepoints. The qPCR signals of BrdU and mouse IgG samples were calculated as percent input values. Fold enrichment of BrdU ChIP over mouse IgG ChIP was calculated for each time point.

II.4.15. Construction and Sequencing of ChIPSeq Libraries:

H3K9me2 and me3 ChIP was performed as above. Five to fifteen nanograms of ChIP DNA or un- enriched whole cell extract (Input) were prepared for sequencing on an Illumina HiSeq2000.

Libraries were constructed with the Truseq DNA sample prep kit V2 (Illumina, CA) with the following modifications: 10ng of ChIP DNA were used as input material. DNA fragments were blunt-ended, 3'-end A-tailed and ligated to indexed TruSeq adaptors. The adaptors were diluted 1:20 to adjust for the input amount of DNA. Indexed adaptors allow for sequencing of multiple samples on the same lane (multiplexing). The adaptor-ligated ChIP DNAs were individually size selected on a 2% agarose gel (Ex-Gel, Life Technologies, CA) to obtain the ligated fragments 300-800bp in length. Size-selected DNAs were amplified by PCR to selectively enrich for those fragments that have adapters on both ends. Amplification was carried out for 15 cycles with the Kapa HiFi polymerase (Kapa Biosystems, Woburn, MA) to reduce the likeliness of multiple
identical reads due to preferential amplification. The final libraries were quantitated by qPCR on an ABI 7900, to allow for accurate quantitation and maximization of number of clusters in the flowcell. Final amplified libraries were also run on Agilent bioanalyzer DNA 7500 LabChips (Agilent, Santa Clara, CA) to determine the average fragment size and to confirm the presence of DNA of the expected size range.

The libraries were pooled and loaded onto a lane of an 8-lane flowcell for cluster formation and sequenced on an Illumina HiSeq2000. The libraries were sequenced from one end of the molecules to a total read length of 100nt. The raw .bcl files were converted into demultiplexed compressed fastq files using Casava 1.8.2.

II.5. Figures

Figure II.1. ORCA interacts with multiple repressive histone lysine methyltransferases.
Figure II.1 (Cont.) (A) a. IP using ORCA Ab from U2OS cells. ORCA, G9a and Suv39H1 were analyzed by immunoblotting (IB). b. IP using G9a Ab from U2OS cells. Endogenous ORCA, G9a and Suv39H1 were analyzed by IB. (B) a and b) Immunoprecipitation (IP) using ORCA antibody (Ab) from cells expressing T7-ORCA and different Flag-KMTs: a. H3K9 KMTs G9a; b. H3K9 KMTs GLP and Suv39H1. (C) U2OS 2-6-3 CLTon cells co-transfected with individual YFP-Lacl-KMTs and CFP-ORCA. Inset represents 150% magnification of the boxed region. (D) IP using T7 ab from cells co-expressing T7-ORCA and; a. Flag-G9a or b. Flag-Suv39H1 in the presence (+) or absence (-) of EtBr. (E) Direct interaction of ORCA and: a. G9a or b. SUV39H1 using purified proteins. ‘*’ denotes cross reacting band and ‘⇒’ denotes ORCA.
**Figure II.2. Interaction of ORCA with histone methyltransferases.**

(A) a. Endogenous ORCA IP in untransfected (UT) U2OS and in U2OS cells expressing full length (FL) or the SET domain of HA-Suv39H1. b. Endogenous ORCA IP in untransfected (UT) U2OS and in U2OS cells expressing full length (FL) or the Ankyrin (ANK) domain of HA-G9a. (B) IB showing efficient depletion of endogenous ORCA from U2OS nuclear extract by using ORCA antibody. (C) Flag IP in Hela cells stably expressing Flag-HA-G9a or Flag-HA-GLP. IPs were conducted using nuclear soluble (S) or chromatin (P) fractions and endogenous ORCA, ORC2, MCM3, Geminin and PCNA were analyzed by IB. (D) ORCA does not interact with arginine methyltransferase (RMT) PRMT5. IP using T7 Ab from cells expressing T7-ORCA and HA-PRMT5.
Figure II.3. ORCA associates with KMT in a chromatin-context dependent manner.  

(A). Schematic representation of various truncation mutants of ORCA containing a T7-epitope on the N-terminus. The specific domains that can associate with G9a and Suv39H1 based on −IB (Figure S2Aa and S2Ab) is depicted as ‘+’. b. Schematic representation of various truncation mutants of G9a containing a HA-epitope on the N-terminus. The interaction domain of G9a that interacts with ORCA (Figure 2Ba) is denoted as ‘+’. c. Schematic representation of various truncation mutants of Suv39H1 containing a HA-epitope on the N-terminus. The interaction domain of Suv39H1 that interacts with ORCA (Figure 2Bb) is denoted as ‘+’.  

(B) a. IP in U2OS cells expressing various HA-G9a mutants and T7-G9a using T7 Ab and analysis by T7 and HA-IB.  b. IP in U2OS cells expressing various HA-Suv39H1 mutants and T7-G9a using T7 Ab and analysis by T7 and HA-IB. ‘***’ denotes the cross-reacting band. (C) a. Cells co-transfected with YFP-LacI (negative control) and CFP-G9a or YFP-LacI-ORCA or the truncation mutants along with CFP G9a in CLTon cells. b. The % of cells with CFP-G9a recruited to the locus is plotted.  

(D) a. Cells co-transfected with YFP-LacI (negative control) and CFP-ORCA or YFP-LacI-G9a and the mutants which are catalytically inactive along with CFP-ORCA in CLTon cells. b. The % of cells with CFP ORCA recruited to the locus.
Figure II.3 (Cont.) (E) a. U2OS 2-6-3 CLTon cells cotransfected with YFP-LacI-ORCA and CFP-G9a in the presence and absence of doxycycline. b. The % of cells with CFP-G9a recruited to the locus in both conditions. Scale bars equal 10 µm. Inset represents 150% magnification of the boxed region. Error bars represent s.d, n=3. ****p<0.0001.
Figure II.4. WD domain of ORCA interacts with H3K9 KMTs
Figure II. 4 (Cont.) (A) a. IP in U2OS cells expressing various T7-ORCA mutants and HA-G9a using T7 Ab and analysis of G9a by IB. b. IP in U2OS cells expressing various T7-ORCA mutants and HA-Suv39H1 using T7 Ab and analysis of Suv39H1 by IB. IP in U2OS cells expressing various HA-G9a mutants and T7-G9a using T7 Ab and analysis by T7 and HA IB. (B) a. Cells were co-transfected CFP-LacI and YFP-G9a (negative control) or CFP-LacI-ORCA along with YFP-G9a truncation mutants. b. The % of cells with YFP-G9a truncation mutants recruited to the locus is plotted. Note the significant reduction in the recruitment of YFP-G9a (aa1-965), the mutant lacking the SET domain, to the locus. Error bars represent s.d, n=3. ****p<0.0001. (C) Localization of H3K9me2 in CLTon cells in cells with (+) or without (-) the expression of YFP-LacI-G9a. Note the H3K9me2 accumulation in YFP-LacI-G9a expressing cells. (D) IP in U2OS cells expressing T7-ORCA and GFP-G9a-full length and ΔSET mutant using T7 Ab and analysis of GFP-G9a by IB. (E) Localization in CLTon cells expressing CFP-LacI-ORCA, of HP1α and YFP-CDK9, at heterochromatic (-Dox) as well as decondensed locus (+Dox). the expression of YFP-LacI-G9a. Note the loss of HP1α and accumulation of YFP-CDK9 upon decondensation of the locus. Scale bar, 10µm. Inset represents 150% magnification of the boxed region. (F) Tethering of YFP-LacI-Orc3 recruits Orc2 at heterochromatic (-Dox) as well as decondensed locus (+Dox). Scale bar, 10µm. Inset represents 150% magnification of the boxed region.
Figure II.5. ORCA and H3K9 KMTs exist in one multimeric complex.

(A) a. Representative single-molecule fluorescence time trajectories for YFP tagged molecules that exhibit one-step, two-step and three-step photobleaching. b. Key to the schematics of the SiMPull assay.
**Figure II.5 (Cont.) (B)** a-b. Schematic and TIRF images of YFP molecules pulled down from U2OS cell lysates expressing T7-ORCA and YFP-ORC1 using biotinylated T7 Ab. The same lysate incubated with biotinylated HA Ab served as the control. c. Average number of YFP fluorescent molecules per imaging area (5000µm²). d. Photobleaching step distribution for YFP-ORC1 bound to T7-ORCA. Note 1:1 ratio of ORCA to Orc1. e. Intensity profiles of the YFP-ORC1 molecules bound to T7-ORCA. (C) a-d. ORCA-G9a pulldown. Shown are YFP molecules pulled down from U2OS cell lysates expressing T7-ORCA and YFP-G9a. Note 1:1 or 1:2 ratio of ORCA to G9a. (D) a-d. ORCA-Suv39H1 pulldown. Shown are YFP molecules pulled down from U2OS cell lysates expressing T7-ORCA and YFP-Suv39H1. Note 1:1 ratio of ORCA to Suv39H1. (E) a-c. Determination of ORCA complexes containing both ORC and G9a by SiMPull and colocalization analyses. a. Schematic of YFP and mCherry molecules pulled down from U2OS cell lysates expressing T7-ORCA, YFP-ORC1 and mCherry-G9a using biotinylated T7 Ab. The same lysate incubated with biotinylated HA Ab served as the control. b. Average number of YFP and mCherry fluorescent molecules per imaging area (5000µm²). c. c. Note 39±5% overlap. Transfection condition used as indicated in Figure 3- Figure Supplement 1Aa, lane3. (F) a-c. Determination of ORCA complexes containing multiple H3K9 KMTs by SiMPull and colocalization analyses. a. Schematic of YFP and mCherry molecules pulled down from U2OS cell lysates expressing T7-ORCA, YFP-Suv39H1 and mCherry-G9a using biotinylated T7 Ab. The same lysate incubated with biotinylated HA Ab served as the control. b. Average number of YFP and mCherry fluorescent molecules per imaging area (5000µm²). c. Note 55±7% colocalization. Transfection condition used as indicated in Figure 3- Figure Supplement 1Ba, lane3. Scale bars, 10µm. Error bars represent s.d, n=3.
Figure II.6. ORC-ORCA-H3K9 KMTs exist in a single complex.

(A) a. Titration of T7-ORCA, mCherry-G9a and YFP-ORC1 plasmids in U2OS cells. b-c. Determination of ORCA complexes containing both ORC and G9a by SiMPull and colocalization analyses. b. Average number of YFP and mCherry fluorescent molecules per imaging area (5000 µm²). c. Note 41±4% overlap. Transfection condition used as indicated in Fig. S3Aa, lane 5.

(B) a. Titration of T7-ORCA, mCherry-G9a and YFP-Suv39H1 plasmids in U2OS cells. b-c. Determination of ORCA complexes containing both G9a and Suv39H1 by SiMPull and colocalization analyses. b. Average number of YFP and mCherry fluorescent molecules per imaging area (5000 µm²). c. Note 46±11% overlap.
Figure II.6 (Cont.) overlap. Transfection condition used as indicated in Fig. S3Ba, lane 5. (C) Sequential IP of HA-Orc1 followed by T7-ORCA from U2OS extracts expressing T7-ORCA, HA-Orc1 and Flag-G9a. IB of G9a corroborated the presence of Orc1-ORCA-G9a triple complex. (D) Sequential IP of HA-G9a followed by T7-ORCA from U2OS extracts expressing T7-ORCA, HA-G9a and Flag-Suv39H1. IB of Suv39H1 corroborated the presence of G9a-ORCA-Suv39H1 triple complex.
Figure II.7. ORCA binds and regulates levels of H3K9 methylation.

(A) a. Schematic of experimental setup for peptide pulldown and analyses by SiMPull. b. TIRF images of YFP-ORCA WT and aa1-127 pulled down by H3K9 modified peptides. Note that the YFP-ORCA WT and aa1-127 truncation mutant expressing lysates were diluted so that the concentration of the overexpressed proteins is comparable (200 and 800 times respectively for WT and aa1-127). c. Average number of fluorescent molecules per imaging area. Scale bars, 10µm. (B) a. Chromatin fractionation in ORCA-depleted U2OS cells followed by IB analysis of H3K9me2 and me3. b. Chromatin fractionation in ORCA depleted diploid fibroblasts, WI38 followed by IB analysis of H3K9me2 and me3. Splicing factor, SRSF1 is shown as a loading control. Error bars represent s.d, n=3. S and S2- cytosolic; S3- nuclear soluble and MNase sensitive; P: nuclear; P3: nuclear insoluble and MNase resistant fraction.
Figure II.8. ORCA binds to H3K9 peptides.

(A) Peptide pulldown using baculovirally purified His-ORCA and N-terminal histone H3 peptides, which are unmodified or acetylated, mono-, di- or tri methylated at K9. (B) a. Schematic of GFP pulldown to quantitate YFP-ORCA expression levels by SiMPull. b. TIRF images of YFP-ORCA WT and aa1-127 pulled down by GFP Ab. Note that the aa1-127 truncation mutant is much more highly expressed compared to WT (5000 fold dilution of aa1-127 shows greater number of molecules/imaging area compared to 2000 fold diluted WT.  c. Quantitation of average number of fluorescent molecules (YFP-ORCA WT and aa1-127) per imaging area in (Bb).
Figure II.9. Loss of ORCA leads to significant reduction in H3K9 methylation.  
(A) Model-based analysis of ChIP-sequencing (MACS) 1.4 peaks analysis of H3K9me3 ChIP-seq in control and ORCA-depleted cells. (B) Regions showing greater than 5-fold decrease in H3K9me3 upon ORCA knockdown plotted along the length of the chromosomes in which they reside. (C) a. Normalized number of reads of repetitive sequences in control and ORCA knockdown H3K9me3 ChIP-seq. Normalized number of reads of b. telomeric repetitive sequences and; c. centromeric repetitive sequences in control and ORCA knockdown H3K9me3 ChIP-seq.
Figure II.9 (Cont.) (D) a-d. Representative regions showing significant decrease in reads in H3K9me3 ChIP on ORCA knockdown compared to the control. (E) a. HA-ORCA ChIP at H3K9me3-target sites and (b) C-FOS. (F) a. Suv39H1 ChIP and b. IgG ChIP at regions showing decrease in H3K9me3. Error bars represent s.d, n=3. C-FOS is shown as negative control.
Figure II.10. ORCA depletion causes changes in chromatin architecture.
Figure II.10 (Cont.) (A) a. Regions with H3K9me3 peaks detected by ChIP-seq plotted along the length of the chromosomes in which they reside. Chromosome scale indicated at the bottom of the chromosomes. b. Regions showing less than 1.3 fold decreases in H3K9me3 upon ORCA knockdown plotted along the length of the chromosomes in which they reside. Chromosome scale is indicated at the bottom of the chromosomes. (B) a-c. Representative regions showing significant decrease in the reads in H3K9me3 ChIP on ORCA knockdown. c-FOS, a region which doesn’t show decrease in H3K9me3 is also shown. (C) a-b. H3K9me2 ChIP at regions showing decrease in H3K9me3. (D) Chromatin fractionation in ORCA depleted U2OS cells and G9a and Suv39H1 IB analyses. SRSF1, a splicing factor, was used as loading control. (E) a. G9a ChIP and b. IgG ChIP at regions showing alterations in H3K9me2 and me3. Error bars represent s.d, n=3.
Figure II.11. ORCA is a scaffold for G9a-Suv39H1 complexes.
**Figure II.11 (Cont.)** (A) a-b. HA-IP in control and ORCA-depleted U2OS cells co-expressing with HA-G9a and Flag-Suv39H1. (B) a. TIRF images of GFP SiMPull in control and ORCA-depleted U2OS cells cotransfected with YFP-Suv39H1 and mCherry-G9a. The same lysates incubated with biotinylated HA Ab served as the control. b. Average number of YFP fluorescent molecules per imaging area (5000µm²). c. The % of mCherry-G9a pulled down by YFP Suv39H1 in control and ORCA knockdown. (C) a. TIRF images of GFP SiMPull in U2OS cells transiently transfected with YFP-Suv39H1, mCherry-G9a and T7-ORCA full-length or truncation mutant 1-270 or 128-647. The same lysates incubated with biotinylated HA Ab served as the control. b. Average number of YFP fluorescent molecules per imaging area (5000µm²). c. The % of mCherry-G9a pulled down by YFP-Suv39H1. The % of mCherry-G9a pulled down by YFP-Suv39H1 in WT-ORCA is 25±1%; 1-270 ORCA is 14±3% and 128-647 ORCA is 29±6%. Scale bars, 20µm. Error bars represent s.d, n=3. **p<0.01, ***p<0.001
Figure II.12. Loss of ORCA causes defects in heterochromatin organization.

(A) IB showing efficient siRNA-mediated knockdown of ORCA. (B) Distribution of S phase cells displaying early, mid and late replication patterns in control and ORCA knockdown cells. Error bars represent s.d, n=3 independent experiments with 500 BrdU positive cells scored in each. (C) BrdU-PI flow cytometry of control and ORCA knockdown cells. (D) Replication timing of genomic regions that show reduced H3K9me3 upon ORCA knockdown. Gray bars represent late replicating domains and black bars denote early replicating domains. HeLa-S3 G1b and HeLa-S3 S1 are late G1 and early S cell cycle fractions that together represent the early replicating regions of the genome.
Figure II.13. ORCA knockdown alters the replication timing.

(A) Patterns of BrdU incorporation in S phase. Examples of early (1), mid (2 and 3) and late (4 and 5) S patterns. Scale bar, 10µm. (B) Replication timing of genomic regions that show reduced H3K9me3 upon ORCA knockdown. Gray bars represent late replicating
Figure II.13 (Cont.) domains and black bars denote early replicating domains. (C) BrdU ChIP in S phase in control and ORCA knockdown cells. Note the changes in replication timing of CELSR3 (b) and FAM20A (c) upon loss of ORCA. C-FOS locus is used as a control region whose replication timing remains unaffected upon loss of ORCA (a). Fold enrichment in the graph represents the % input of BrdU ChIP over % input of rIgG ChIP.
Figure II.14. Heterochromatin organization role of ORCA is independent of its role in preRC assembly.

(A) Schematic of depletion of ORCA using the proteotuner system. (B) Western blotting showing the levels of endogenous and exogenous ORCA in the presence of control and ORCA siRNA. β”, a nuclear speckle protein, serves as the loading control. Note that the DD-T7-ORCA-siRNA NTV is stabilized upon the addition of Shield1. (C) Chromatin fractionation and immunoblotting showing the levels of chromatin bound Orc2 in control and ORCA siRNA-treated cells (either in the absence or presence of exogenous ORCA). Note the reduction in chromatin bound Orc2 in the absence of ORCA and the rescue of its levels upon expression of exogenous ORCA. Also note the increase in the soluble pool of Orc2 in the absence of ORCA and the decrease of its levels upon expression of exogenous ORCA. Splicing factor, SRSF1 is shown as a loading control. (D) IB showing the levels of endogenous and exogenous ORCA at G1/S and 12h post release from aphidicolin. H3 is used as loading control.
Figure II.14 (Cont.) (E) a-b. Patterns of BrdU incorporation in control and ORCA depleted cells in late S phase. The white arrowheads indicate preferential incorporation of BrdU incorporation at perinucleolar regions upon loss of ORCA. Scale bar, 10µm. b. % increase in S phase cells displaying early and % decrease of the mid and late replication patterns in ORCA-depleted cells compared to control cells. Error bars represent s.d, n=3 independent experiments with ~450 BrdU positive cells scored in each.(F) a. H3K9me3 and HP1α immunofluorescence in control and ORCA depleted cells. The white arrowheads indicate H3K9me3 and HP1α immunofluorescence at perinucleolar regions upon loss of ORCA. Representative regions in control and ORCA-depleted cells marked by white dotted squares (1, 2 and 3) are shown at 3X magnification on the right. Scale bar, 10µm. b. The % of cells with HP1α at nucleolar periphery in control and ORCA-depleted cells. Error bars represent s.d, n=3.
Figure II.15. BrdU incorporation preferentially at perinucleolar regions in cells lacking ORCA.
Figure II.15 (Cont.) BrdU incorporation preferentially at perinucleolar regions in cells lacking ORCA.

(A) IP of DD-T7-ORCA siRNA NTV from U2OS cells using T7 Ab. DD-T7-ORCA siRNA NTV and endogenous Orc were analyzed by IB. (B) Patterns of BrdU incorporation in control and ORCA depleted cells in late S phase. Scale bar, 10µm. (C) H3K9me3 and HP1α immunofluorescence in control and ORCA depleted cells. Scale bar, 10µm. (D) Flow cytometry of control and ORCA knockdown cells at 0, 4, 8 and 12h post release from Aphidicolin block.
Figure II.16. Model depicting the role of ORCA in organizing heterochromatin

Model representing mode of regulation of heterochromatin by ORCA.
II.6 References


Jenuwein. 2012. Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. Cell. 150:948-960.


CHAPTER III. UNDERSTANDING THE ROLE OF ORCA IN HETEROCHROMATIN ORGANIZATION AND REPLICATION – INTERACTION WITH H4K20 LYSINE METHYLTRANSFERASES

III.1. Introduction

Every cycling cell needs to duplicate its genetic material accurately and then segregate the chromosomes faithfully to the daughter nuclei. Furthermore, the epigenetic information must also be restored from one cell generation to the next. Proteins involved in DNA replication play crucial roles in the inheritance of chromatin domains. (Chakraborty et al., 2011b; Li and Zhang, 2012) However, the mechanism that ensures that chromatin architecture is reestablished once the replication has been accomplished remains to be elucidated and is an intense area of research. (Abmayr and Workman, 2012; Giri and Prasanth, 2012) Several studies in budding yeast, fission yeast, Drosophila and humans have pointed out that DNA replication proteins coordinate heterochromatin organization and gene silencing either by facilitating nucleosome assembly of heterochromatin or by recruiting factors that are key to the establishment and maintenance of heterochromatin or by coordinating with the siRNA machinery to maintain heterochromatin. (Li and Zhang, 2012) Whether the role of replication initiation factors in heterochromatin assembly is independent of their replication initiation function remained to be determined.

Initiation of DNA replication in eukaryotes requires the sequential assembly of a multiprotein complex at the origins of replication. (Bell and Dutta, 2002) Origin Recognition complex (ORC) consisting of six subunits serves as the landing pad for the establishment of pre-replication complex during G1 phase of the cell cycle. (Bell and Stillman, 1992) ORC-Associated (ORCA)/LRWD1 is an ORC interacting protein that is required for stabilizing ORC onto chromatin. (Shen et al., 2010) In diploid fibroblasts, depletion of ORCA causes an accumulation of cells in G1 phase of the cell cycle. (Shen et al., 2012) In addition, previous work from our laboratory has shown that ORCA regulates

Chapter III has been published with few modifications as: Giri, S., and Prasanth, S.G., 2015. Association of ORCA/LRWD1 with repressive histone methyltransferases mediates heterochromatin organization. Nucleus.
replication initiation by modulating the interaction between pre-replicative complex component Cdt1 and its inhibitor Geminin.(Shen et al., 2012)

While the role of ORC in replication initiation has been extensively studied and is relatively well understood, the role of ORC and ORCA in heterochromatin remained unclear. Others and we have previously shown that ORCA binds to heterochromatin, including at centromeres and telomeres in human and mouse cells.(Bartke et al., 2010; Chan and Zhang, 2012; Shen et al., 2010; Vermeulen et al., 2010) By using Stable Isotope Labelling in Cell Culture (SILAC) and modified N-terminal histone tails as baits, it was shown that ORCA binds specifically to repressive trimethylated H3K9, K27 and H4K20 marks.(Vermeulen et al., 2010) We recently conducted a study to investigate the function of ORCA at H3K9me3-containing heterochromatic domains.(Giri et al., 2015) We found that ORCA interacts with multiple H3K9 KMTs G9a/GLP and Suv39H1. A multimeric complex containing all the H3K9 KMTs including G9a, GLP, Suv39h1 and SETDB1 is recruited to pericentric heterochromatin and aids in maintenance of H3K9me2 and me3.(Fritsch et al., 2010) By using Single Molecule Pulldown assays we found that ORC-ORCA-H3K9 KMTs exist in a single complex. The existence of a complex containing ORC and H3K9 KMTs is very exciting as it reiterates the importance of a cross-talk between eukaryotic DNA replication proteins and the repressive epigenetic machinery. ORCA also directly binds to H3K9me2 and me3 with stronger binding to the trimethylated mark. The loss of ORCA resulted in the global reduction of H3K9me3, consistent with the observation that loss of ORCA also showed reduced association of Suv39H1 on chromatin. In addition to the reduction in the levels of H3K9me3, there was also a reduction of H3K9me2 upon depletion of ORCA. By using chromatin immunoprecipitation we further showed that ORCA binds to specific regions on the chromatin that are enriched for H3K9me2 and me3. Furthermore, the H3K9me2 and 3 are lost specifically from these regions when ORCA was depleted from human cells. In order to understand the role of ORCA in the complex containing G9a and Suv39H1, we investigated the stability of these complexes upon the loss of ORCA. We found that the loss of ORCA resulted in the reduction of the complexes containing G9a and Suv39H1. Furthermore, aberrant chromatin organization also resulted in defective DNA replication in cells depleted of ORCA. Specifically, we found that loss of ORCA showed a reduction
of late replication patterns and aberrant replication timing. This was specifically at the regions that showed loss of H3K9me2 and me3. To tease out the role of ORCA in DNA replication initiation versus chromatin organization, we depleted ORCA at the G1/S transition and then probed for defects in DNA replication. We found that postG1 cells showed aberrant chromatin organization as evident by the mislocalization of HP1α and H3K9me3. Our results showed that ORCA regulates chromatin organization independent of its role in DNA replication initiation.

III.2. ORCA interacts with multiple repressive methyltransferases

Our work has demonstrated that ORCA interacts with repressive H3K9 KMTs G9a/GLP and Suv39H1. Since ORCA has been shown to bind to repressive histone lysine methylation marks, specifically H3K9me3, H3K27me3 and H4K20me3,(Bartke et al., 2010; Chan and Zhang, 2012; Giri et al., 2015; Vermeulen et al., 2010) we investigated whether ORCA also associates with the KMTs that catalyze H4K20 and H3K27 methylation. To investigate this we utilized U2OS 2-6-3 CLTon cells.(Chakraborty et al., 2014; Janicki et al., 2004) This cell line has a tandem array of Lac operator (LacO) sequences inserted into a single locus in the cells. The locus is heterochromatic and can be visualized by the stable expression of mCherry-Lac repressor (LacI) in the cells. As observed earlier, tethering of H3K9 KMT G9a to the locus showed robust accumulation of ORCA at this site (Fig. III. 1A). In order to examine whether ORCA interacts with H3K27 KMT EZH2, we tethered YFP-LacI-EZH2 to the locus and found that CFP-ORCA is also recruited to the locus (Fig. III. 1A). This is interesting because a recent report shows that there is a functional cross-talk between H3K9 and H3K27 KMTs in mouse embryonic stem cells.(Mozzetta et al., 2014) H3K9 KMTs were found to cooperate with the H3K27 methylation complex, the Polycomb Repressive Complex 2 (PRC2).(Mozzetta et al., 2015b) G9a and GLP double knockouts showed reduction in PRC2 recruitment and H3K27me3 levels. In addition, the authors showed that G9a monomethylates H3K27 and this aids PRC2-dependent trimethylation of H3K27. It would be interesting to investigate whether ORCA functions in a complex that contains both G9a and EZH2 and whether it regulates the function of this complex. Further, how would such a multimeric complex be recruited to chromatin sites? In addition, it would
be important to investigate whether loss of ORCA leads to upregulation of PRC2-regulated genes. Previous work has shown that loss of ORCA in mouse cells leads to upregulation of centromeric transcription.(Chan and Zhang, 2012) It would be crucial to determine whether this transcriptional control by ORCA is restricted to repetitive heterochromatic regions or whether it extends to repressed euchromatic regions controlled by G9a/GLP and/or PRC2.

H4K20me1 and H4K20me2 are shown to be involved in DNA replication, whereas H4K20me3 is required for pericentric heterochromatin organization.(Jorgensen et al., 2013) Suv4-20H1 and H2 catalyze the di- and tri methylation of H4K20.(Schotta et al., 2004) We tethered YFP-LacI-Suv420H1.1 to the CLTon locus and found that CFP-ORCA is recruited to the locus indicating that ORCA also interacts with Suv420H1.1 (Fig. III. 1A). The presence of H3K9me3 is essential for H4K20 trimethylation at pericentric heterochromatin.(Fischle et al., 2005; Lachner et al., 2001; Stewart et al., 2005)

The mode of recruitment of different KMTs to specific target sites in mammalian cells is not clearly understood.(Mozzetta et al., 2015a) None of the KMTs, except PRDM family members, can bind to DNA directly. The association of KMTs to specific DNA-binding proteins/chromatin binding proteins may provide a means by which they could be targeted to specific sites. The replication protein ORCA associates with specific chromatin marks and the histone modifying machinery via its WD domain. It remains to be determined if ORCA binds to DNA and the DNA modifying machinery directly.

Post-translational modifications on histones are preserved at specific genomic regions from one cell generation to another. It is largely believed that during DNA replication the modified histones present on the parental DNA are randomly segregated to the daughter strands and this is required for the further addition of modifications on the newly assembled histones. Post-translational modifications are transmitted with the parental histones to the newly formed DNA strand.(Alabert et al., 2015) Interestingly, H3K9me3 and H3K27me3 were found to propagate by modification of the parental and newly generated histones and this was found to extend over several cell generations.(Alabert et al., 2015) Recent evidence from Drosophila has pointed out that
the histone modifications, specifically H3K4me3 and H3K27me3, are lost during DNA replication and that histone-modifying machinery associates with specific genomic loci that persists during DNA replication and enables the re-establishment of the histone modifications. (Petruk et al., 2012) During DNA replication, the DNA methyl transferase (DNMT1) and the H3K9 KMT G9a physically and functionally cooperate at the replication fork to coordinate DNA and histone methylation. (Esteve et al., 2006) The association of ORCA with the histone modifying machinery brings many interesting questions: is this association occurring at specific stages of the cell cycle, is it occurring at specific origins, if this association is disrupted can this affect replication timing?

**III.3. The WD-40 domain of ORCA interacts with Suv420H2**

WD repeat-containing proteins bind to histone and nucleosomes and function in a diverse array of cellular functions. (Suganuma et al., 2008) ORCA consists of leucine rich repeats at its N-terminus and a WD-repeat domain at its C-terminus. (Shen et al., 2010) We have found that the WD domain of ORCA associates with the modified histones and also mediates the association to H3K9 KMTs. (Giri et al., 2015) In order to map the interaction of ORCA with H4K20 KMT-Suv420H2, we cotransfected HA-Suv420H2 and T7-ORCA in human U2OS cells. This was followed by immunoprecipitation with the T7 antibody (Fig. III. 2B). ORCA interacted robustly with Suv420H2 (Fig. III. 2Ab). This interaction could be indirect or mediated by multiple domains of Suv420H2. Different truncation mutants of ORCA including 1-127a (Spanning LRR), 1-270aa (LRR+linker), 128-647 (linker+WD) and 270-647aa (WD alone) were co-transfected with HA-Suv420H2 (Fig. 2Aa). Immunoprecipitation using T7 antibody revealed that the WD domain of ORCA mediates the association of ORCA to Suv420H2 (Fig. III. 2Ab). It is interesting to note that the Linker+WD mutant shows reduced association to Suv420H2 compared to the WD alone. Determination of the structure of ORCA would provide insights into the organization of these domains.

Reciprocal immunoprecipitation from cells expressing HA-Suv420H2 and T7-ORCA using HA antibody showed the association of Suv420H2 with ORCA (Fig. III. 2Bb). To further determine the domain of Suv420H2 that interacts with ORCA, we made two truncation mutants of Suv420H2 (Fig. III. 2Ba), the N terminal fragment (1-250aa)
containing the SET domain of Suv420H2 and the C terminal fragment (250-462aa) containing the region required for heterochromatinization. Co-transfection of these HA tagged mutants with T7-ORCA followed by HA immunoprecipitation showed that both the N and C terminal fragments independently interact with ORCA (Fig. III. 2B, lanes 1 and 2). In the future we will investigate whether the interaction between ORCA and Suv420H2 is direct by using purified proteins. Furthermore, we would also determine if this interaction is DNA-dependent.

ORCA was also found to interact with H4K20 monomethylase, PR-SET7 (data not shown). These observations that ORCA interacts with H4K20 KMTs raises many interesting possibilities. Work from Reinberg’s lab (Beck et al., 2012) pointed towards the possibility that during DNA replication initiation Orc1 and ORCA bind to H4K20me2 and me3 respectively, thereby establishing the origins poised for replication in S-phase. While this is an intriguing possibility, it would be crucial to conclusively determine whether this binding of ORC/ORCA to methylated H4K20 is related to the function of these proteins in DNA replication initiation as opposed to chromatin organization in repressive environments.

Based on our results, we speculate that ORCA acts similar to HP1α and facilitates the establishment and maintenance of H3K9 methylation-containing heterochromatin. ORCA acts as a scaffold to hold together the H3K9 KMTs megacomplex and stabilizes them on chromatin. This leads to the establishment of methylated H3K9, thereby providing more binding sites for ORCA and this process continues with the end result of establishment of heterochromatin domains. It is quite possible that ORCA’s binding to methylated H4K20 and interaction with PR-SET7 and Suv4-20H1/2 could be geared towards the exact same purpose at heterochromatin and could be independent of its role in DNA replication initiation. Another hypothesis is that ORCA regulates specific subsets of origins, namely late replicating ones that reside within heterochromatin. In that case, ORCA could have multiple, interdependent functions in heterochromatin, which can be broadly classified into those required for heterochromatin replication initiation and its organization. Determining the ORCA binding sites on the genome would provide important insights into its role in replication initiation and heterochromatin function. This would allow us to
investigate whether ORCA associates with all the origins or predominantly with the later firing ones.

There is also increasing evidence for a functional cross-talk between multiple KMT complexes. For example, the G9a/GLP and the PRC2 interact physically as well as functionally. (Mozzetta et al., 2014) Further, the activity of G9a dictates the recruitment of PRC2 to specific target genes. (Mozzetta et al., 2014) Such a functional cross-talk may exist between multiple histone modifying complexes and further studies would be critical to determine mechanistic details of this interaction.

III.4. ORCA could act as a scaffolding factor in multiple repressive environments

We have previously shown that ORCA associates with constitutive heterochromatin present at centromeres and telomeres. (Shen et al., 2010) In a genome wide RNAi screen to identify factors involved in Xi silencing, ORCA along with ORC was found to be involved in the maintenance of X-chromosome inactivation. (Chan et al., 2011) ORCA was also uncovered in the Xist interactome as a Xist interacting protein. (Minajigi et al., 2015) Since ORCA interacts with EZH2 and associates with H3K27me3, (Vermeulen et al., 2010) this could be the mechanism of association of ORCA with the inactive X chromosome and perhaps Xist and thereby facilitating silencing. In this context it is interesting to note that Orc2 localizes on Xi and impacts Xi silencing. (Chan et al., 2011) It is likely that ORCA functions in the same pathway as Orc2 in mediating Xi silencing. Based on our previous data (Giri et al., 2015) and data shown here we hypothesize that ORCA functions as a scaffolding factor for H3K9, H4K20 and H3K27 methyltransferases at multiple repressive environments such as centromeres, telomeres and inactive X, to name a few (Fig. III. 3). ORCA seems to be crucial for constitutive as well as facultative heterochromatin organization. Constitutive heterochromatin represents gene-poor pericentromeric regions enriched with H3K9me3 and H4K20me3. (Saksouk et al., 2015) Our data supports the model that ORCA acts as a scaffold that enables the KMTs to carry out their function. Since facultative heterochromatin represents genomic regions that can adopt open or closed conformations depending on temporal and spatial contexts, (Trojer and Reinberg, 2007) it would be
crucial to understand the role of ORCA in establishing or maintaining the facultative heterochromatin.

SILAC based proteomic studies has also revealed that ORCA/ORC can bind to methylated CpG DNA. (Bartke et al., 2010) The authors showed that ORCA/ORC bind to methylated DNA and histones in a cooperative fashion. It is interesting to note that the tethering of the repressive KMTs to the CLT on locus enhances ORCA’s interaction with the locus in multiple ways. First, ORCA localizes to the locus by interacting with the KMTs. Second, ORCA can also bind to the methylated H3K9, K27 and H4K20 that are established by the KMTs. Third, the presence of methylated histones could aid ORCA’s binding to methylated DNA. A natural segue to this would be investigating whether ORCA interacts with DNA methyltransferases. Another avenue of investigation that would be exciting to pursue in the light of this data would be to examine the possible role of ORCA in silencing of LINEs and SINEs. Finally, since bivalent (H3K27 me3 and H3K4me3 marked) (Bernstein et al., 2006) and trivalent (H3K9me3, H3K27 me3 and H3K4me3 marked) domains (Bernstein et al., 2006) exist in embryonic stem cells and regulate key events during differentiation, it would be important to investigate the role of ORCA during development.

III.5. Perspectives

ORCA is turning out to be a multifaceted protein playing key roles in heterochromatin organization and DNA replication. Determining genome wide association of ORCA will be crucial for gaining a better understanding of its function. In addition, it will be important to determine the full complement of histone modifications to which ORCA interacts with. Such studies will provide crucial insights into the possible role of ORCA in different chromatin environments. Finally, in the light of its association with EZH2, it will be exciting to investigate the function of ORCA in embryonic stem cells and in the context of development and differentiation.
III.6. Figures

Figure III.1. ORCA interacts with H3K9, K27 and H4K20 histone lysine methyltransferases.

(A) U2OS 2-6-3 CLTon cells co-transfected with individual YFP-LacI-KMTs and CFP-ORCA. Inset represents 200% magnification of the boxed region.
Figure III.2. ORCA interacts with Suv420H2.

(A) a. Schematic representation of various truncation mutants of ORCA containing a T7-epitope on the N-terminus. The specific domains that can associate with Suv420H2 are depicted as ‘+’. IP in U2OS cells expressing various T7-ORCA mutants and HA-Suv420H2 using T7 Ab and analysis by T7 and HA IB.  

(B) a. Schematic representation of various truncation mutants of Suv420H2 containing a HA-epitope on the N-terminus. The interaction domains of Suv420H2 that interact with ORCA are denoted as ‘+’. b. IP in U2OS cells expressing various HA-Suv39H1 mutants and T7-G9a using HA Ab and analysis by T7 and HA IB.
Figure III.3. Model depicting the role of ORCA in organizing different chromatin domains.
**Figure III.3 (Cont.)** Model representing the mode of regulation of chromatin at constitutive and facultative heterochromatin and at origins by ORCA. Inactive X chromosome is depicted as an example of facultative heterochromatin.
III.7. References


CHAPTER IV. UNDERSTANDING THE ROLE OF Orc5 IN CHROMATIN DECONDENSATION

IV.1. Introduction

In eukaryotes, the initiation of DNA replication requires the coordinated action of a multiprotein pre-replication complex at the origins (Bell and Dutta, 2002). The Origin Recognition Complex (ORC), a six-subunit complex, binds to replication origins during G1 phase of the cell cycle and this is followed by a sequential assembly of other preRC components (Bell and Stillman, 1992). In addition to its role in replication initiation, ORC subunits contribute to other cellular processes including transcriptional silencing, heterochromatin organization, sister chromatid cohesion, centrosome duplication, telomere maintenance and cytokinesis (Sasaki and Gilbert, 2007).

Open chromatin structures are known to regulate the efficiency of preRC formation thereby facilitating replication initiation (Papior et al., 2012). However, the molecular mechanisms that affect chromatin structure and how the preRC components establish themselves on the chromatin remain to be understood. The accessibility of the replication factors is influenced by the chromatin structure and the chromatin architecture dictates the efficiency of origin usage and firing (Brown et al., 1991; Ferguson and Fangman, 1992; Simpson, 1990; Stevenson and Gottschling, 1999). Histone acetylation is known to play a key role in the regulation of origins of DNA replication in yeast and Drosophila and there is accumulating evidence that the deacetylation of histone negatively affects origin activity (Aggarwal and Calvi, 2004; Groth et al., 2007a; Knott et al., 2009b; Unnikrishnan et al., 2010; Vogelauer et al., 2002). Furthermore, the replication timing of the beta-globin domain in human cells is also modulated by histone modifications at the origin (Goren et al., 2008). There is accumulating evidence that histone acetyl transferases act as positive regulators of replication origins in yeast, Drosophila as well as human cells (Groth et al., 2007b; Knott et al., 2009a). In yeast GCN5p, a histone acetyl transferase was found to positively stimulate DNA replication by nulling the inhibitory effect of the histone deacetylases (Espinosa et al., 2010;
Vogelauer et al., 2002). Further, Hat1p and its partner Hat2p interact with ORC (Suter et al., 2007). In Drosophila, the HATs Chameau (Chm) and CBP (Nejire) stimulate origin activity (Aggarwal and Calvi, 2004; McConnell et al., 2012). In human cells, HBO1, another HAT, associates with ORC and is required for the loading of MCM onto chromatin and for replication fork progression (Iizuka et al., 2006; Iizuka and Stillman, 1999; Miotto and Struhl, 2008; Miotto and Struhl, 2010). A recent study has pointed out that the acetylation of some histone lysines depends on the binding of ORC to the origin and that the acetylation is at its maximum on the nucleosomes adjacent to one side of the major initiation site (Liu et al., 2012). How ORC regulates such chromatin modifications and how the chromatin structure at origins is organized remains to be defined.

ORC consists of six subunits, and in human cells they are highly dynamic. The largest subunit, Orc1 is degraded at the end of G1 and its rebinding to chromatin is an obligatory step for the establishment of preRC in G1 (Mendez et al., 2002). The smallest subunit of ORC, Orc6, binds to the ORC in a transient manner and also has independent roles in cytokinesis (Bernal and Venkitaraman, 2011; Prasanth et al., 2002). Orc2, 3, 4 and 5 in human cells constitute the core ORC and are associated with each other throughout the cell cycle (Dhar et al., 2001; Vashee et al., 2001). Orc1, 4 and 5 are members of the AAA+ family of ATPases and contain consensus motifs. Mutations in the ATP-binding sites on Orc4 and 5 impair complex assembly, while the ATP binding of Orc1 is dispensable (Ranjan and Gossen, 2006; Siddiqui and Stillman, 2007). Orc2 and 3 also have an AAA+ structure but do not possess a consensus ATP binding motif.

Multiple subunits of human ORC, including Orc1, 2, 3 and 5 and the ORC-associated (ORCA) have roles in heterochromatin organization (Giri et al., 2015; Prasanth et al., 2010; Shen et al., 2010). In yeast, Orc5 has separable functions in replication initiation and silencing (Dillin and Rine, 1997). Further, in Drosophila and humans the loss of multiple ORC subunits leads to chromosome segregation defects (Pflumm and Botchan, 2001; Prasanth et al., 2004). In this manuscript, we report that Orc5 has a distinct function in chromatin unfolding. Ectopic tethering of Orc5 to a chromatin locus leads to dramatic chromatin decondensation. This chromatin-opening role of Orc5 required the activity of the HAT GCN5. The binding of GCN5 to origins is reduced at origins in cells lacking Orc5. We propose that Orc5 subunit of ORC plays a key role in mediating large-
scale chromatin opening that in turn facilitates the loading of other preRC components to the origins.

IV.2. Results

IV.2.1. Ectopic tethering of Orc5 induces large-scale chromatin decondensation

To investigate the chromatin changes that occur when preRC proteins including Origin Recognition Complex (ORC) proteins bind to origins, we tethered individual subunits of ORC to a heterochromatic locus using an in vivo cell system (CLTon) generated in the human U2OS osteosarcoma cells (Fig. IV.1A). This reporter carries a stably integrated 200-copy transgene array with the Lac operator repeats and this heterochromatic locus is visualized by the stable expression of Cherry-Lac repressor (Cherry-LacI). Upon transcriptional activation of this reporter locus, by the addition of doxycycline, this locus shows chromatin decondensation (Janicki et al., 2004; Shen et al., 2010). We generated triple-fusion proteins of YFP-LacI-ORCs and these were tethered to the CLTon locus. Targeting YFP-LacI to this locus showed association of LacI to the heterochromatic CLTon locus (Fig. IV.1Ba). Surprisingly, tethering of YFP-LacI-Orc5 caused dramatic decondensation at the CLTon locus whereas none of the other ORC subunits including Orc1, Orc2, Orc3, Orc4 and Orc6 caused any changes to the chromatin architecture at the locus (Fig. IV.1Ba). 81% of YFP-LacI-ORC5-tethered cells showed decondensation of the heterochromatic locus (Fig. IV.1Bb). Furthermore, the extent of decondensation upon tethering Orc5 to the locus was determined by calculating the area of the decondensed chromatin. Measurement of the area of decondensation upon tethering Orc5 revealed a range of chromatin decondensation ranging from 2-35 µm² (Fig. IV.1Bc), whereas the control YFP-LacI showed condensed loci with size in the range of 0.2-1.3 µm² (Fig. IV.1Bc). Based on this, we categorized the Orc5 mediated decondensation phenotype into 3 categories: medium (2-6 µm²), large (6-10 µm²) and very large (10-35 µm²) (Fig. IV.1C). The tethering of Orc5 to the locus showed 37%, 34% and 29% of cells showing medium, large and very large range of decondensation respectively.

We investigated the role of Orc5 in chromatin decondensation by utilizing another system, in this case a CHO-derived A03 cell line that contains 90Mb of homogenously staining region generated through stable integration and amplification of the LacO-DHFR
vector (Li et al., 1998). Tethering Orc5 to the A03 locus also showed dramatic decondensation of this locus (Fig. IV.1D). The decondensation upon tethering Orc5 was in the range of 4.5-27 µm² whereas tethering of YFP-LacI showed decondensation in the range of 0.6-1.2 µm².

We next determined the minimum domain of Orc5 that is required for its chromatin decondensation ability. Triple-fusion-Orc5 truncation mutants (including YFP-LacI-Orc5.1-100aa; 101-200aa; 201-300aa; 301-400aa and 301-435aa) were generated (Fig. IV.2A) and their ability to cause chromatin unfolding was examined (Fig. IV.2B). As described earlier, YFP-LacI-Orc5 FL caused decondensation of 81% of the CLTon locus (Fig. IV.2C). The Orc5 truncation mutants 1-100, 101-200, 201-300 and 301-400aa failed to show chromatin unfolding, but the N-terminal truncation 301-435aa mutant showed chromatin unfolding in ~40% of cells (Fig. IV.2B and C). The decondensation upon tethering YFP-LacI-Orc5 (301-400aa) was in the range of 0.5-1.0 µm² whereas tethering of YFP-LacI-Orc5 (301-435aa) showed decondensation in the range of 2.5-19.0 µm² (Fig. IV.2D). Our results indicated that the last 35 amino acids at the C-terminus of Orc5 are critical for its chromatin decondensation function.

Upon examination of the C-terminal 400-435aa (Fig. IV.3A), we observed that it was enriched with acidic residues. It has previously been reported that targeting of acidic activators to heterochromatic chromatin domains can cause large-scale chromatin decondensation (Carpenter et al., 2005). We generated a mutant of Orc5 where multiple aspartic acid residues were replaced by alanines (Fig. IV.3A). However, this mutant, when tethered to the locus, showed similar levels of chromatin decondensation, suggesting that the ‘acidic-domain’ within Orc5 is not required for decondensation (Fig. IV.3B). We next determined if the ATP binding ability of Orc5 is required for its chromatin-unfolding function. We generated Walker A mutant (K43A) and an arginine finger mutant (R166A) (Fig. IV.3A) and tethered these to the CLTon locus (Fig. IV.3C). The extent of chromatin decondensation upon tethering these mutants was comparable to that of the wild type Orc5 suggesting that the ATP binding ability of Orc5 is also dispensable for its chromatin decondensation function.
IV.2.2. Orc5 associates with the histone acetyl transferase GCN5

Histone acetylation, catalyzed by various histone acetyl transferases (HATs) is linked with the open chromatin state and this is known to facilitate transcription (Narlikar et al., 2002). Since Orc5 was found to cause chromatin decondensation, we addressed its association with known histone acetyl transferases, namely GCN5. We co-transfected T7-Orc5 and Flag-GCN5 and carried out immunoprecipitation with Flag antibody. Orc5 and GCN5 were found to interact with one another (Fig. IV.4A). A reverse immunoprecipitation experiment recapitulated the interaction (Fig. IV.4B T7-Orc5 FL lane).

We next mapped the region within Orc5 that associates with GCN5. We generated several truncation mutants of Orc5 (Fig. IV.4C) and co-transfected each of these with Flag-GCN5. Immunoprecipitation with T7 antibody revealed that full-length Orc5 and the 100-435aa fragment efficiently associates with GCN5 (Fig. IV.4B), suggesting that the first 100 aa within the N-terminus of Orc5 are dispensable for GCN5 binding. Orc2 was found to bind to full-length Orc5 and to fragments 100-435 and 300-435 (Fig. IV.4B and 4C). These results suggest that Orc2 and GCN5 could associate with Orc5 simultaneously and the binding may not be mutually exclusive. Further, tethering of GCN5 to the CLTon locus also showed robust recruitment of Orc5 to the site, corroborating the immunoblot results (Fig. IV.4D).

We next determined if GCN5 could itself cause chromatin decondensation in our CLTon assay. YFP-LacI-GCN5 was tethered to the locus and the status of chromatin architecture at the CLTon locus was evaluated (Fig. IV.4Ea). GCN5 could also mediate chromatin decondensation, however this decondensation looked visually different from what was observed for Orc5. While tethering YFP-LacI-Orc5 caused considerably large-scale decondensation, tethering YFP-LacI-GCN5 caused a smaller scale ‘puffy’ appearance of chromatin, but still significant decondensation when compared to YFP-LacI control (Fig. IV.4Da). While the range of chromatin decondensation upon tethering YFP-LacI-Orc5 varied from 2-35 µm², the extent was much smaller for YFP-LacI-GCN5, which showed decondensation in the range of 3-9.5 µm² (Fig. IV.4Eb).
IV.2.3. Orc5-mediated chromatin decondensation is GCN5-dependent

To gain an insight into the functional relevance of the Orc5 and GCN5 interaction, we evaluated if GCN5 is required for the Orc5-mediated chromatin decondensation. We depleted GCN5 using siRNA in CLTon cells and examined if tethering Orc5 could still induce chromatin decondensation. Remarkably, in GCN5-depleted cells, we observed a significant decrease (32% decrease, ****p <0.0001) in the extent of Orc5-mediated chromatin decondensation (Fig. IV.5A-C). In control cells, 84.3% of the cells showed chromatin decondensation upon tethering Orc5. On the other hand, this number reduced to 52% upon loss of GCN5 (Fig. IV.5A). To better understand this reduction in Orc5 mediated chromatin decondensation, we scored CLTon cells based on medium or large decondensation of the heterochromatic locus. Upon loss of GCN5, there was a striking reduction in cells showing large decondensation (from 44% in control cells to 16% in GCN5 knockdown cells) (Fig. IV.5B). This result was corroborated by examining the area of the loci in control and GCN5-depleted cells. The area of CLTon locus varied from 2 – 35 µm² in control cells and 1 – 7 µm² in GCN5-depleted cells (Fig. IV.5C).

We next examined the status of various chromatin marks at the Orc5-tethered locus. Immunofluorescence using H3K9me3 antibody showed robust accumulation of this mark at the LacI-containing heterochromatic locus (Fig. IV.6A). However, upon tethering Orc5 to the locus, H3K9me3 was distinctly devoid at these sites (Fig. IV.6A). Since Orc5 associates with GCN5, we examined if histone H3 acetylation was accumulating in Orc5-tethered cells. We did not observe robust accumulation of H3acetyl marks in the highly decondensed Orc5-tethered cells; however, the control cells were clearly devoid of this mark at the condensed CLTon locus (Fig. IV.6B).

IV.2.4. Loss of Orc5 causes reduction in histone acetylation at origins

Histone acetylation is required for origin activation during S-phase (Unnikrishnan et al., 2010). It has also been shown to regulate the timing of replication origin firing (Vogelauer et al., 2002). We asked if Orc5-mediated chromatin opening facilitates histone acetylation in cooperation with GCN5 at the origins of replication. We conducted H3Ac ChIP in control and Orc5-depleted cells and found small but reproducible reduction in the H3 acetylation at select origins but not at distal sites of the specific origin.
(Fig. IV.6C). These results suggest that Orc5 might make the chromatin more accessible for the establishment of preRCs during G1 at specific origins (Fig. IV.7). Our results provide newer insights into the specific role of one of the ORC subunits in facilitating chromatin opening for preRC establishment.

IV.3. Discussion

Replication of DNA occurs once and only once per cell division cycle. The licensing of replication origins requires the sequential binding of ORC, Cdc6 and Cdt1 that is needed for the loading of the Mcm2-7 complex to the chromatin (Bell and Dutta, 2002). ORC, consisting of six subunits, serves as the landing pad for the assembly of this multiprotein complex at the origins of replication. The contribution of individual subunits in this process remains to be understood. We demonstrate that Orc5 subunit is unique and when tethered ectopically to a transgene array, induces large-scale chromatin decondensation. It associates with the histone acetyl transferase, GCN5, an H3 histone acetylase. Orc5’s ability to cause chromatin decondensation requires GCN5 (Fig. 7) and the loss of Orc5 causes decreased acetylation at specific origins.

GCN5 is a global regulator of gene expression (Baker and Grant, 2007; Robert et al., 2004). More recently, GCN5 has been found to associate with yeast origins, albeit weakly, (Espinosa et al., 2010) and positively regulates DNA replication by counteracting the inhibitory effects of HDACs. In addition, GCN5-mediated acetylation of H3 lysines has also been proposed to function in replication-coupled nucleosome assembly (Burgess et al., 2010). We propose that Orc5 at origins helps the recruitment of GCN5 to these sites and this in turn facilitates the opening of chromatin, thus enabling the loading of other preRC components.

The acetylation of histone H3 and H4 is known to be dynamically regulated around the origins of replication that facilitate origin firing (Unnikrishnan et al., 2010). Furthermore, H4K79ac is enriched at origins; H4K16ac is enriched at early firing origins and also limits the spread of heterochromatin at origins. H3K9/14ac and H4K5, 8, 12 ac are enriched at active origins and are believed to promote firing (Dorn and Cook, 2011).

The HAT, Hbo1, is an H4-specific histone acetyl transferase that interacts with human ORC and MCMs (Iizuka and Stillman, 1999). It is required for replication licensing
(Iizuka et al., 2006) and is known to associate with replication origins (Miotto and Struhl, 2008). It has recently been demonstrated that hyperacetylation of histone H4 mediated by Hbo1, increases Mcm2-7 loading to chromatin (Miotto and Struhl, 2010). It is generally believed that histone acetylation loosens up the compacted chromatin, thereby increasing the accessibility to load MCM helicase complex or act as a molecular tag to which the helicase is tethered (Chadha and Blow, 2010). Hbo1 association to origins is dependent on Cdt1, and it has been shown that Cdt1 can modulate chromatin accessibility through temporal recruitment of Hbo1 to origins (Miotto and Struhl, 2008). Interestingly, ectopic tethering of Cdt1 is also known to induce large-scale chromatin unfolding at a transgene array (Wong et al., 2010). It is well established that the origins consist of open chromatin during G1 and then become less accessible as cells exit out of G1 phase (Djeliova et al., 2002; Pemov et al., 1998). The replication origins are known to exhibit temporal dynamics in chromatin structure, with highly open structure during G1 and more closed architecture during S phase. Elegant work has shown that endogenous replication origins including Mcm4 and Lamin display more open chromatin structure during G1 than in S-phase using qPCR approach on DNase-1 treated chromatin samples (Wong et al., 2010).

In addition, tethering of replication protein Cdc45 leads to chromatin decondensation in a Cdk2 dependent fashion (Alexandrow and Hamlin, 2005). Such decondensation is mediated by H1 phosphorylation. The phosphorylation of H1, in turn, is Cdk2-dependent and leads to chromatin decondensation in S phase. These observations point towards a role for Cdc45 in replication fork progression by large-scale chromatin changes (Alexandrow and Hamlin, 2005). Our observations on the CLTon locus demonstrating that Orc5 can efficiently cause chromatin decondensation provide strong indication that similar events occur at endogenous origins, whereby Orc5 could cause local chromatin changes in collaboration with GCN5. Another piece of evidence to support this idea comes from a recent study, which showed preferential association of Orc5 with H3K27Ac (Ji et al., 2015) as compared to several other activating marks such as H3K4me3, H3K79me2 and H3K36me3.

Orc5 is one of the ORC subunits and ATP binding to Orc5 is involved in efficient ORC formation (Siddiqui and Stillman, 2007; Takahashi et al., 2004). Orc5 has also been
implicated in silencing at the HML and MHR loci in yeast and in heterochromatin organization in human cells (Dillin and Rine, 1997; Prasanth et al., 2010). The gene encoding Orc5 maps to chromosome 7q22 and is frequently deleted in adult acute myeloid leukemia, myelodysplastic syndrome, uterine leiomyomas and malignant myeloid diseases (Frohling et al., 2001; Quintana et al., 1998). We have observed the dose-dependent chromatin decondensation at the CLTon locus and this is directly correlated with the expression level of Orc5. This is supported by the fact that tethering of Orc5 to the CLTon locus can cause dramatic chromatin decondensation. However, tethering of other ORC subunits does not, despite the fact that Orc5 can be recruited to the locus by other ORC subunits. Our data implies that the excessive levels of Orc5 in the cell could result in aberrant chromatin decondensation and cause genomic instability.

IV.4. Materials and methods:

pEGFP-LacI vector was a kind gift from Dr. Miroslav Dundr (Kaiser et al., 2008) and used to pEYFP-LacI vector. YFP-LacI-Orc1 through 6 were cloned by amplifying and inserting the Orcs into pEYFP-LacI vector. YFP-LacI-Orc5 D414,426,433A, YFP-LacI-Orc5 K43A, YFP-LacI-Orc5 R166A were generated by site directed mutagenesis (Stratagene) with YFP-LacI-Orc5 wildtype as template. T7 Orc5 full length and truncation were cloned by amplifying and inserting into pCGT vector. Flag-GCN5 was kind gifts of Dr. Brian Freeman at UIUC.

IV.4.1. Cell culture:

U2OS Osteosarcoma cells were grown in high glucose Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal bovine serum (FBS – Hyclone). U2OS -2-6-3 CLTon cells were cultured in DMEM supplemented with 10% Tet system approved FBS (Clonetech).

IV.4.2. Immunofluorescence and fluorescent protein visualization:

For visualizing YFP-LacI tagged proteins, cells were fixed with 2% Formaldehyde in phosphate buffered saline (PBS -pH 7.4) for 15 minutes in room temperature followed by permeabilization with 0.5% Triton X-100 in PBS for 7 minutes on ice followed by blocking. For H3K9me3 and H3Ac immunofluorescence, cells were pre-extracted before
fixing with 0.5% Triton X-100 in Cytoskeletal buffer (CSK: 100mM NaCl, 300mM Sucrose, 3mM MgCl₂, 10mM PIPES at pH 6.8) for 5 minutes on ice followed by fixing with 1% Formaldehyde in phosphate buffered saline (PBS - pH 7.4) for 5 minutes in room temperature and then blocking. This was followed by blocking for 30 minutes with 1% Normal goat serum (NGS) in PBS. After that, 1 hour of Primary antibody incubation in a humidified chamber followed by secondary antibody incubation for 25 minutes was then carried out. Nuclei were then stained with DAPI and mounted using vectashield (Vector Laboratories Inc.). The following antibodies were used for immunofluorescence: H3K9me3 (1:200, Millipore 07-523) and H3 Ac (1:500, Millipore 06-599).

IV.4.3. siRNA mediated depletion of Orc5:

For Orc5 knockdown, U2OS cells were grown to 30% confluency followed by 2 rounds of knockdown with control luciferase gene (Shen et al., 2010) and Orc5 si (Prasanth et al., 2010), 24h apart at a final concentration of 100nM using Lipofectamine RNAiMax (Invitrogen). This was followed by collection of cells for ChIP 24h after the second knockdown.

For GCN5 knockdown, CLTon cells were grown to 30% confluency on coverslips followed by 2 rounds of knockdown with control luciferase gene (Shen et al., 2010) and GCN5 si (Palhan et al., 2005) 24h apart at a final concentration of 40nM. YFP-LacI-Orc5 was transfected while carrying out the second round of knockdown and the cells were fixed for microscopy 24h later. The first round of knockdown was carried out using Lipofectamine RNAiMax (Invitrogen) and the second round of knockdown along with transfection was carried out using Lipofectamine 2000 (Invitrogen).

IV.4.4. Immunoprecipitations and immunoblots:

For co-immunoprecipitation experiments, co-transfections were carried out in U2OS cells. Flag-HATs and T7-Orc5 were transiently transfected and cells were lysed, 24 hours post-transfection, in IP buffer (50mM HEPES pH 7.9, 10% Glycerol, 200mM NaCl, 0.1% Triton X-100, 1mM CaCl₂) supplemented with the protease and phosphatase inhibitors. 4U of MNase (Sigma) was then added per 10 cm plate followed by nutation at room temperature for 20 min. EDTA (final concentration 5mM) was added to stop the reaction followed by centrifugation at 12500 rpm, 5 min at 4°C. The supernatant was
used for pre-clearing. Pre-clearing was carried out with Gammabind Sepharose beads for 1 hour and the lysates were incubated with appropriate antibody overnight. For pulldown of the antibody bound complexes, agarose beads were washed in the same IP buffer and were incubated with lysate containing antibodies for 2h. This was followed by 3 washes of the pulled down complexes and finally denaturation of the pulled down proteins by the addition of Laemmli buffer and incubation on heatblock (4°C) for 10 min. The complexes were then analyzed by Western blotting.

For immunoprecipitations and immunoblotting the following antibodies were used anti-Flag M2 (1:500, Sigma), anti-T7 (1:5000; Novagen), anti-ORC2 pAb 205-6 (1:1000).

IV.4.5. Chromatin Immunoprecipitation

Control and Orc5 knockdown cells were crosslinked for 10 mins at room temperature by addition of Formaldehyde (Sigma) to culture medium to a final concentration of 1%. The reaction was stopped by addition of glycine at a final concentration of 0.125 M. Fixed cells were washed twice (quickly) in chilled PBS and harvested with PBS. Chromatin was prepared by two subsequent extraction steps (10 min at 4°C) with Buffer 1 (50 mM Hapes/KOH pH 7.5; 140 mM NaCl; 1 mM EDTA; 10% Glycerol; 0.5% NP-40; 0.25% Triton) and Buffer 2 (200 mM NaCl; 1mM EDTA; 0.5mM EGTA; 10 mM Tris pH 8). Nuclei were then pelleted by centrifugation, resuspended in SDS Lysis Buffer (50 mM Tris pH 8; 0.1% SDS; 1% NP-40; 0.1% Na-Deoxycholate; 10 mM EDTA; 150 mM NaCl) and subjected to sonication with Bioruptor Power-up (Diagenode) for 45 min (3 times of 15 min sonication – with each cycle being 30” On and 30” Off). This yielded genomic DNA fragments 300 bp in length. The obtained Chromatin was then precleared with Protein A/G ultralink beads (53133, Pierce) for 1h at 4°C and immunoprecipitation with the H3 Ac antibody and rabbit IgG was carried out overnight at 4°C. Immune complexes were pulled down by adding pre-blocked protein A/G ultralink beads and incubated for 2 h at room temperature. Beads with immune complexes bound to them were washed once with Low salt buffer (0.1% SDS; 1% Triton; 2 mM EDTA; 20 mM Tris pH 8; 150 mM NaCl), once with High salt buffer (0.1% SDS; 1% Triton; 2 mM EDTA; 20 mM Tris pH 8; 500 mM NaCl), once with LiCl wash buffer (10 mM Tris pH 8.0; 1% Na-deoxycholate; 1% NP- 40, 250 mM LiCl; 1 mM EDTA) and twice with TE.
Beads were then eluted twice (10 min each) in TE + 1% SDS + 0.1% NaHCO₃ at 65°C and the cross-links were then reversed O/N at 65°C. DNA from the eluted material was then treated with RNase (10 µg/ml) for 1h at 37°C followed by 2h of Proteinase K treatment (4ul 0.5M EDTA, 8ul 1M Tris pH 6.9, 1ul Proteinase K 20mg/ml) at 42°C. DNA was isolated by using QIAquick PCR purification kit (Qiagen) resuspended in elution buffer and q-PCR performed using PowerSYBR Green PCR Master mix (Applied Biosystems) and analyzed on a 7300 PCR System (Applied Biosystems). ChIP-qPCR results were analyzed and plotted as percentage (%) of IP/input signal (% input).

The primer sequences of the regions analyzed are as follows:

- MCM4, −5 kbp forward, TTCACATCCACCCAGCTTATC
- MCM4, −5 kbp reverse, AGAGCATTCTTCCCCTGATG
- MCM4 origin, reverse, TTGGGTGGCTACTTGGTGTT
- MCM4 origin, reverse, TAGGCCCCTCGCTTTGTT
Figure IV.5. Figures

Figure IV.1. Orc5 causes chromatin decondensation.
**Figure IV.1 (Cont.)** (A) Schematic of the heterochromatic locus in U2OS 2-6-3 CLTon cells. (B) a. Chromatin decondensation upon tethering YFP-LacI and YFP-LacI-Orc1 through 5 to the heterochromatic locus of CLTon cells. Inset represents 200% magnification of the boxed region. Scale bar, 10µm. b. The % of cells with open loci upon tethering either YFP-LacI or YFP-LacI-Orc1 through 5 to the heterochromatic locus of CLTon cells. Error bars represent s.d, n=3. c. Area of heterochromatic loci upon tethering YFP-LacI, and YFP-LacI-Orc5. Error bars represent s.d, n=3. ****p<0.0001. (C) Chromatin decondensation upon tethering YFP-LacI and YFP-LacI-Orc5 to the heterochromatic locus of CLTon cells. Inset represents 200% magnification of the boxed region. Scale bar, 10µm. (D) Chromatin decondensation upon tethering YFP-LacI and YFP-LacI-Orc1 through 5 to the heterochromatic locus of AO3 cells. Inset represents 200% magnification of the boxed region. Scale bar, 10µm.
Figure IV.2. The last 35 amino acids of Orc5 are necessary for its function in chromatin decondensation.

(A) Chromatin decondensation upon tethering YFP-LacI and various truncation mutants of YFP-LacI-Orc5 to the heterochromatic locus of CLTon cells. Inset represents 200% magnification of the boxed region. Scale bar, 10µm. (B) Schematic representation of various truncation mutants of Orc5 containing a T7-epitope on the N-terminus. The specific domains that can cause decondensation are indicated with ‘+’.

(C) The % of cells with decondensed chromatin in the presence of different YFP-LacI-Orc5 truncation mutants.

(D) Area of loci with decondensed chromatin for different YFP-LacI-Orc5 truncation mutants.
Figure IV.2 (Cont.) with open loci upon tethering either YFP-LacI or various truncation mutants of YFP-LacI-Orc5 to the heterochromatic locus of CLTon cells. Error bars represent s.d, n=3. (D) Area of heterochromatic loci upon tethering YFP-LacI or YFP-LacI-Orc5 truncation mutants. Error bars represent s.d, n=3.
Figure IV.3. Decondensation by Orc5 is independent of its acidic residues and its ATPase activity.

(A) Schematic representation of the acidic activation domain of Orc5. The Aspartates mutated to Alanine are marked by ‘*’. (B) Chromatin decondensation upon tethering YFP-LacI, YFP-LacI-Orc5 and YFP-LacI-Orc5 D414,426,433A to the heterochromatic locus of CLTon cells. Inset represents 200% magnification of the boxed region. Scale bar, 10µm. (C) Chromatin decondensation upon tethering YFP-LacI, YFP-LacI-Orc5 and YFP-LacI-Orc5 K43A and YFP-LacI-Orc5 R166A to the heterochromatic locus of CLTon cells. Inset represents 200% magnification of the boxed region. Scale bar equals 10µm.
Figure IV.4. Orc5 interacts with GCN5.
(A) IP in U2OS cells expressing T7-Orc5 and Flag-GCN5 using Flag Ab and analysis by T7 and Flag immunoblot (IB). (B) IP in U2OS cells expressing various T7-Orc5 mutants
Figure IV.4 (Cont.) and Flag-GCN5 using T7 Ab and analysis by T7, Flag and Orc2 IB. (C) Schematic representation of various truncation mutants of Orc5 containing a T7-epitope on the N-terminus. The specific domains that can associate with GCN5 and Orc2 based on IB (Figure 4B) is depicted as ‘+’.

(D) Tethering of YFP-LacI-GCN5 to CLTon locus shows recruitment of Orc5 to the site. (E) a. Chromatin decondensation upon tethering YFP-LacI, YFP-LacI-Orc5 and YFP-LacI-GCN5 to the heterochromatic locus of CLTon cells. Inset represents 200% magnification of the boxed region. Scale bar, 10µm. b. Area of heterochromatic loci upon tethering YFP-LacI, YFP-LacI-Orc5 and YFP-LacI-GCN5. Error bars represent s.d, n=3.
Figure IV.5. Orc5 causes decondensation in a GCN5 dependent fashion.
(A) The % of YFP-LacI-Orc5 tethered cells with open loci in control and GCN5 knockdown cells. Error bars represent s.d, n=3. (B) Extent of decondensation upon tethering YFP-Lacl-Orc5 in control and GCN5 knockdown cells. Error bars represent s.d, n=3. (C) Area of heterochromatic loci upon tethering YFP-Lacl and YFP-LacI-Orc5 (in control and GCN5 knockdown cells). Error bars represent s.d, n=3. ****p<0.0001.
Figure IV.6. Orc5 mediated chromatin decondensation is accompanied by the loss of H3K9me3.

(A) Localization of H3K9me3 in CLTon cells upon tethering YFP-LacI and YFP-LacI-Orc5. Note the loss of H3K9me3 in YFP-LacI-Orc5 expressing cells. Scale bar, 10µm.
(B) Localization of H3Ac in CLTon cells upon tethering YFP-LacI and YFP-LacI-Orc5. Note the exclusion of H3Ac upon tethering YFP-LacI and no obvious accumulation upon tethering YFP-LacI-Orc5. Representative regions in YFP-LacI and YFP-LacI-Orc5 tethered cells marked by white dotted squares (1, 2 and 3) are shown at 4X magnification below. Scale bar, 10µm. (C) H3Ac ChIP at MCM4 origin and -4kb region in U2OS cells after control and Orc5 knockdown. Error bars represent s.d, n=3.
Figure IV.7. Cartoon depicting Orc5-GCN5 mediated chromatin opening.
IV.6. References


Li, G., G. Sudlow, and A.S. Belmont. 1998. Interphase cell cycle dynamics of a late-replicating, heterochromatic homogeneously staining region: precise choreography of


CHAPTER V. CONCLUSIONS AND PERSPECTIVES

The Origin Recognition Complex (ORC) has a myriad of functions in a eukaryotic cell. Apart from its indispensable role in replication initiation, ORC has roles in cohesion of sister chromatids, cytokinesis and neurogenesis (Sasaki and Gilbert, 2007). In addition, ORC has a well-documented role on heterochromatin. It associates with Hp1α staining heterochromatic foci and is required for the integrity of heterochromatin (Prasanth et al., 2010). While close to two decades worth of work on ORC’s association with heterochromatin exist, surprisingly little is known about ORC’s role at these repressive chromatin environments. ORCA - identified by our lab couple of years back as an ORC interacting protein, also associates with heterochromatin (Bartke et al., 2010; Chan and Zhang, 2012; Shen et al., 2010; Vermeulen et al., 2010). With that in mind, I began by asking the following questions: “why do ORCA and ORC associate with heterochromatin?” and “what is the functional significance of this association?”.

As I describe in Chapter II, I find that ORCA interacts with multiple H3K9 lysine methyltransferases (KMTs) like G9a, GLP and Suv39H1. Interestingly, while ORCA can interact with G9a and Suv39H1 directly, its interaction with G9a on chromatin requires G9a’s catalytic activity. In addition, activation of transcription at a heterochromatic locus abolishes this interaction at the locus. These data together indicated that ORCA’s interaction with G9a at a heterochromatic locus requires the catalytic activity of G9a.

Also, that the interaction is robust at a transcriptionally silent region. ORC, ORCA, G9a and Suv39H1 exist in a single complex. Loss of ORCA leads to loss of H3K9me2 in both cancerous U2OS cells and primary diploid fibroblast WI38. ChIP-seq in U2OS cells upon loss of ORCA shows a global decrease in H3K9me3 with close to 18% of the H3K9me3 peaks showing greater that 5-fold decrease upon loss of ORCA. These regions also show decrease in H3K9me2 by ChIP-qPCR. ORCA ChIP-qPCR in U2OS shows strong binding of ORCA to these regions indicating that ORCA binds and regulates the regions that show strong decrease in H3K9me3 upon loss of ORCA. Suv39H1 ChIP-qPCR upon loss of ORCA shows a reduction of Suv39H1’s association with chromatin indicating that this might be the reason why we observe H3K9me3 decrease upon loss of ORCA. These changes in chromatin architecture upon loss of ORCA translate into defects in late replication and changes in replication timing of heterochromatic loci. In addition, I find
that these defects in chromatin organization and late replication are replication initiation
independent functions of ORCA. I find that massive heterochromatin organization
defects happen by knocking down ORCA at G1/S stage of cell cycle, even though this is
a time point by which origins have been properly licensed.

While I was investigating the interaction of ORCA with repressive H3K9 KMTs, I also
found ORCA to be interacting with H3K27 KMT EZH2 and H4K20 KMTs PR-SET7
and Suv420H1 and H2. As I discuss in chapter III, these are extremely interesting pieces
of data as they point to ORCA having multiple role at repressive chromatin environment.
In light of ORCA’s interaction with EZH2, it would be interesting to investigate the role
of ORCA in the context of differentiation and X chromosome inactivation. Since
previous work from our lab has shown that ORCA is always in a complex with Orc2
(Shen and Prasanth, 2012), at the very least, these investigations will also reveal novel
functions of Orc2 and most likely other Orcs. Additionally, ORCA’s interaction with the
H4K20 KMTs could point to additional functions in replication licensing and
pericentromeric and telomeric organization.

It would now be crucial to determine ORCA’s binding to the genome and to compare it
with ORC binding. This would throw light on the regions where ORC exists as a complex
with ORCA. As I show in Chapter II, a subset of ORCA-ORC complexes exists at late
replicating heterochromatin. Now it is imperative to understand whether ORCA marks all
origins along with ORC, or whether it marks a certain subset of origins, like the late
replicating ones. Carrying out ORCA ChIP-seq would provide information on the
genome wide occupancy of ORCA. This, in conjugation with ORC ChIP-seq would
reveal the regions of the genome bound by both ORCA and ORC. In addition, carrying
out ORC ChIP-seq upon conditions of ORCA knockdown will provide information about
the regions of genome where ORCA regulates ORC’s association with chromatin.

While Chapters II and III focus on ORCA and ORC’s role in establishing and
maintaining heterochromatin, Chapter IV deals with an opposing function of a
component of ORC, Orc5. I find that Orc5 causes chromatin decondensation and
positively regulates the levels of H3 acetylation at origins of DNA replication. I find that
tethering Orc5 to the heterochromatic CLTon locus causes large scale decondensation of
the locus. By making smaller truncations of the protein I find that the last 35 amino acids of Orc5 are crucial for this decondensation effect. In order to understand the mechanism of the decondensation process, I investigate Orc5’s interaction with histone acetyltransferases and find that Orc5 robustly interacts with the HAT GCN5. Knocking down GCN5 causes a significant decrease in the extent of decondensation. In addition, upon tethering Orc5, the heterochromatic locus is maximally decondensed in G1. This points to the possibility that Orc5 causes an opening up of chromatin at G1 phase of cell cycle and this can be happening at origins of DNA replication since these are regions where ORC resides. Interestingly, tethering Orc5 to the locus causes the replication timing of the locus to change from late to early S. In this light, it would be interesting to investigate whether Orc5 is a global regulator of replication timing in mammalian cells.

Knocking down Orc5 causes an increase in G1 stage of cell cycle and a concomitant decrease in S phase by both flow cytometry and MCM-PCNA immunofluorescence. In addition, loss of Orc5 causes reduction in H3 acetylation at a subset of origins. Acetylation of H3 was one of the earliest identified marks of origins in higher eukaryotes. How acetylation is established at origins is not well understood. The ORC associated HAT, HboI, is thought to be involved in the process. H3 acetylation is thought to provide a permissive chromatin environment for origin firing. My data together indicated that Orc5 causes a permissive chromatin environment and aids in H3 acetylation at origins. This is required for origin establishment and firing. So in the absence of Orc5 there is decrease in S phase and G1 arrest.

This role of Orc5 opens up an extremely interesting line of investigation. As I show in Chapter V, one process that is affected upon loss of Orc5 is DNA replication. Another process that could be affected is transcription. It would be crucial to look at H3Ac levels by ChIP-seq upon loss of Orc5. It would tell us whether only origins are affected or whether acetylation of promoters and enhancers is also affected. It would also be important to look at the H3K27 acetylation at super enhancers upon loss of Orc5. Superenhancers contain clusters of enhancers with lineage specific functions (Hnisz et al., 2013; Parker et al., 2013; Whyte et al., 2013). Cancerous cells acquire superenhancers that regulate key proliferation genes (Chapuy et al., 2013; Groschel et al., 2014; Loven et al., 2013; Mansour et al., 2014; Northcott et al., 2014). It would be interesting to see if
loss of Orc5 causes decrease in acetylation of these regions. Loss of Orc5 causes proliferation defects in cancerous U2OS cells. This could be a reflection of inefficient origin firing upon loss of Orc5. In addition to that, it could also be due to inefficient transcription of key proliferative genes. So H3K27 ChIP-seq along with RNA-seq of Orc5 depleted cells will provide key insights into the physiological significance of the decondensation caused by Orc5.
V.1. References


